Renormalization-group approach to interacting fermions

R. Shankar

Sloane Laboratory of Physics, Yale University, New Haven, Connecticut 06520

The stability or lack thereof of nonrelativistic fermionic systems to interactions is studied within the renormalization-group (RG) framework, in close analogy with the study of critical phenomena using ϕ^4 scalar field theory. A brief introduction to ϕ^4 theory in four dimensions and the path-integral formulation for fermions is given before turning to the problem at hand. As for the latter, the following procedure is used. First, the modes on either side of the Fermi surface within a cutoff Λ are chosen for study, in analogy with the modes near the origin in ϕ^4 theory, and a path integral is written to describe them. Next, an RG transformation that eliminates a part of these modes, but preserves the action of the noninteracting system, is identified. Finally the possible perturbations of this free-field fixed point are classified as relevant, irrelevant or marginal. A $d=1$ warmup calculation involving a system of fermions shows how, in contrast to mean-field theory, which predicts a charge-density wave for arbitrarily weak repulsion, and superconductivity for arbitrarily weak attraction, the renormalization-group approach correctly yields a scale-invariant system (Luttinger liquid) by taking into account both instabilities. Application of the renormalization group in $d=2$ and 3, for rotationally invariant Fermi surfaces, *automatically* leads to Landau's Fermi-liquid theory, which appears as a fixed point characterized by an effective mass and a Landau function F, with the only relevant perturbations being of the superconducting (BCS) type. The functional flow equations for the BCS couplings are derived and separated into an infinite number of flows, one for each angular momentum. It is shown that similar results hold for rotationally noninvariant (but time-reversal-invariant) Fermi surfaces also, with obvious loss of rotational invariance in the parametrization of the fixed-point interactions. A study of a nested Fermi surface shows an additional relevant flow leading to charge-density-wave formation. It is pointed out that, for small Λ/K_F, a $1/N$ expansion emerges, with $N=K_F/\Lambda$, which explains why one is able to solve the narrow-cutoff theory. The search for non-Fermi liquids in $d=2$ using the RG is discussed. Bringing a variety of phenomena (Landau theory, charge-density waves, BCS instability, nesting, etc.) under the one unifying principle of the RG not only allows us to better understand and unify them, but also paves the way for generalizations and extensions. The article is pedagogical in nature and is expected to be accessible to any serious graduate student. On the other hand, its survey of the vast literature is mostly limited to the RG approach.

CONTENTS

I. INTRODUCTION

This article is an expanded version of a short paper (Shankar, 1991) in which the application of the renormalization-group (RG) methods to interacting nonrelativistic fermions in more than one spatial dimension was considered. It contains more technical details than its predecessor and is much more pedagogical in tone. Several related topics are reviewed here so that readers with a variety of backgrounds may find the article accessible and self-contained. Consequently each reader is likely to run into some familiar topics. When this happens he or she should go through the section quickly to ensure that this is indeed the case and get used to the no-

 0034-6861/94/66(1)/129(64)/$11.40

tation and conventions. Upon reading this article, the readers with a condensed-matter background will see how the RG allows us to synthesize a variety of seemingly unrelated phenomena in condensed-matter theory, such as Landau's Fermi-liquid theory, the BCS instability, charge-density-wave and spin-density-wave instabilities, nesting, and so on. Readers familiar with the RG, but not these topics, will see that, by following a route parallel to the one that led to a very successful treatment of critical phenomena, we are automatically led to many known results in the above-mentioned topics and newer ways of understanding them. However, there are also many fascinating differences between critical phenomena and the phenomena considered here, which make this approach very interesting from the point of view of the RG. At the time of writing, there are relatively few new results, and the emphasis is on a deeper understanding and unification of the older results the RG affords us. However, the machinery developed here, especially for anisotropic systems, has the potential for changing this state of affairs in the not too distant future. The author is working on a few new applications and hopes the readers will find many more.

The concept of the RG was first introduced by Stuckelberg and Petermann (1953). Its implications for quantum electrodynamics were explored in a seminal paper by Gell-Mann and Low (1954). These concepts were extended and generalized by Callan (1970) and Symanzik (1970).

What is the RG? When we speak of a group in quantum mechanics we are thinking of symmetry operations, i.e., transformations that leave the physics invariant. What is the transformation here? Let us consider quantum electrodynamics. When we compute a physical quantity like the scattering rate between electrons, in a power series in the coupling constant α, we find that the coefficients of the series are given by integrals over particle momenta k, and that these in turn diverge because the allowed values of k go up to infinity. These *ultraviolet divergences* are at variance with experiment, which gives finite answers for all physical quantities. Renormalization is the way to reconcile these two facts. In this scheme, one first cuts off all integrals at the *cutoff* Λ. This gives answers that are finite, but dependent on Λ, which is an artifact in continuum theory. To get around this, one asks if it is possible to choose for each cutoff Λ a corresponding coupling $\alpha(\Lambda)$ so that the physical quantities like scattering amplitudes come out Λ independent. (In quantum electrodynamics one must also renormalize the mass of the electron with the cutoff. The word coupling shall mean all such parameters that define the theory.) It is by no means obvious that this can be done in every field theory. However, in the case of quantum electrodynamics or any *renormalizable* field theory, one can prove that, to any given order in perturbation theory, it is possible to choose a handful of parameters of the model in a cutoff-dependent way so as to make physics at momenta much smaller than the cutoff independent of it. Since the cutoff is eventually sent to infinity, this means at any finite momentum. *This change in the cutoff by a factor s, accompanied by a suitable change in couplings, is an invariance of the theory.* These transformations form a group with the composition rule that a change by a factor s_1 followed by a change by a factor s_2 should equal a change by a factor $s_1 s_2$. If we write

$$s = e^{-t} \tag{1}$$

so that

$$\Lambda(t) = \Lambda_0 e^{-t} \tag{2}$$

where Λ_0 is some fixed number, the group composition law is that when two transformations are implemented in sequence the parameters t add.

A central quantity in this approach is the *β function,* defined as

$$\beta(g) = \frac{dg}{dt} , \tag{3}$$

where g is the generic name for the coupling constant(s). Our convention is the one used in condensed-matter physics, wherein increasing t decreases the cutoff. The field theorists use the opposite convention and differ by a sign. To avoid all confusion let us consider the case of Yang-Mills theory, in which

$$\frac{dg}{dt} = cg^3 + \text{higher orders}, \quad c > 0 . \tag{4}$$

If we integrate this equation from $t = 0$ to $t = t$ (so that the cutoff changes from Λ_0 to $\Lambda_0 e^{-t}$), we find

$$g^2(t) = \frac{g^2(0)}{1 - 2g^2(0)ct} . \tag{5}$$

What this equation tells us is that, as we send the cutoff to infinity (t to $-\infty$), we must reduce the coupling to zero logarithmically:

$$g^2(t) \simeq 1/|t| . \tag{6}$$

The point to notice in all of the above is that one is interested in how to vary the cutoff only with the intention of eventually sending it to infinity, which is where it belongs in a continuum theory. Had the theory been free of ultraviolet divergences, the question of changing the couplings and the cutoff, keeping the physics invariant, might never have come up. It is clear that from this vantage point the RG has no place in condensed-matter physics, where the degrees of freedom live on a lattice and there is a natural cutoff on all momenta: $\Lambda \simeq 1/a$, where a is the lattice constant. (For a notable exception, see C. DiCastro and G. Jona Lasinio, 1969.)

This point of view was dramatically altered following the work of Kadanoff (1965) and Wilson (1971), who gave a different and more physical interpretation of renormalization. In this modern view one contemplates changing the cutoff (and the couplings) even in a problem where nature provides a natural cutoff, such as the inverse lattice spacing, and there are no ultraviolet infinities. We

shall now discuss an example from statistical mechanics in which the value of such a procedure is apparent. The discussion will be schematic, since a more detailed one follows in the next section.

Let us consider a cubic lattice (in d dimensions) with a real scalar field $\phi(\mathbf{n})$ at each site labeled by the vector $\mathbf{n}$ with integer coefficients. The classical statistical mechanics of this system is described by the partition function

$$Z=\int \prod_{\mathbf{n}} d\phi(\mathbf{n}) e^{S(\phi(\mathbf{n}))} . \tag{7}$$

This is just the usual sums over configurations with the Boltzmann factor $e^{-\beta\varepsilon}$ written in terms of the *action S,* also called the *Hamiltonian.* (Both terms will be used interchangeably to prepare the reader for what happens all the time in the literature.) As long as the number of sites is finite, S is just a regular function and Z is just a multiple integral. In the limit of infinite sites, S becomes a functional and Z becomes a *functional integral* or a *Feynman path integral.* Feynman introduced his path integral to describe quantum-mechanical problems in d spatial dimensions as a sum over classical configurations in $d+1$ dimensions. Thus our Z could very well stand for Feynman's representation of a quantum problem in one lower dimension, and the following considerations apply to it.[1] For problems with bosonic operators, the derivation of the path integral can be found in Sec. III.B of the review by Kogut (1979). The derivation for the fermionic problem will be given in Sec. III.

A typical quantity one is interested in is the average of the correlation between the variables at two different sites, also called the *two-point function:*

$$G(\mathbf{n}_1,\mathbf{n}_2)=G(\mathbf{n}_1-\mathbf{n}_2) \text{ (assuming translation invariance)} \tag{8}$$

$$\equiv\langle\phi(\mathbf{n}_1)\phi(\mathbf{n}_2)\rangle \tag{9}$$

$$=\frac{\int \prod_{\mathbf{n}} d\phi(\mathbf{n})\phi(\mathbf{n}_1)\phi(\mathbf{n}_2)e^{S(\phi(\mathbf{n}))}}{\int \prod_{\mathbf{n}} d\phi(\mathbf{n})e^{S(\phi(\mathbf{n}))}} . \tag{10}$$

For long separations this *correlation function* typically falls off exponentially as

$$G(\mathbf{n}_1-\mathbf{n}_2)\simeq e^{-|\mathbf{n}_1-\mathbf{n}_2|/\xi} , \tag{11}$$

where ξ is the *correlation length.*[2] The exception is when the parameters are such that the system is at a *critical point,* as in the case of a magnet undergoing a Curie transition from the ferromagnetic to the paramagnetic state. In this case it falls like a power

$$G(\mathbf{n}_1-\mathbf{n}_2)\simeq\frac{1}{|\mathbf{n}_1-\mathbf{n}_2|^x} , \tag{12}$$

where x is a *critical exponent.* Other critical exponents characterize other power laws at the critical point. A remarkable feature, which we shall address shortly, is that several systems with microscopically distinct Hamiltonians (or actions) have the same critical exponent.

In the case of quantum problems written as path integrals, the correlation length is related to m, the *mass gap,* or the lowest excitation energy above the ground state as per

$$\xi=1/m , \tag{13}$$

and the critical case corresponds to $m=0$.

A. Three stages of RG transformation

An equally complete description of the above system is possible in terms of the Fourier transforms,

$$\phi(\mathbf{k})=\frac{1}{V}\sum_{\mathbf{n}} e^{i\mathbf{k}-\mathbf{n}}\phi(\mathbf{n}) , \tag{14}$$

where V is the volume of the system. The allowed momenta $\mathbf{k}$ lie within a Brillouin cube of sides $2\pi/a$ in all directions. The partition function becomes

$$Z=\int \prod_{|\mathbf{k}|\le\pi/a} d\phi(\mathbf{k})e^{S(\phi(\mathbf{k}))} , \tag{15}$$

and $G(\mathbf{k})$, the Fourier transform of $G(\mathbf{n}_1-\mathbf{n}_2)$, is given by

$$\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle=(2\pi)^d\delta^{(d)}(\mathbf{k}_1+\mathbf{k}_2)G(\mathbf{k}_1) \tag{16}$$

$$=\frac{\int \prod_{\mathbf{k}} d\phi(\mathbf{k})\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)e^{S(\phi(\mathbf{k}))}}{\int \prod_{\mathbf{k}} d\phi(\mathbf{k})e^{S(\phi(\mathbf{k}))}} . \tag{17}$$

Let us now imagine that we are interested only in the physics at long distances (compared to the lattice spacing a), for example, in $G(\mathbf{r})$ for large separations $\mathbf{r}$. In momentum space this translates to small $\mathbf{k}$. To be specific let us say we are interested only in correlations of modes that lie within a tiny ball of size Λ/s (with s very large) centered at the origin, and not interested in the modes that lie in the rest of the Brillouin zone, taken to be a sphere (rather than a cube) of radius $2\pi/a$. (This modification makes no difference to the small-k asymptotics.) We shall refer to the small-k modes as "slow modes" and the others as "fast modes." Let us define two sets of variables:

$$\phi_<=\phi(k) \text{ for } 0<k<\Lambda/s \text{ (slow modes)} , \tag{18}$$

$$\phi_>=\phi(k) \text{ for } \Lambda/s\le k\le\Lambda \text{ (fast modes)} . \tag{19}$$

[1]Dimension $d+1$ is of course time. One works with imaginary time, with the option to continue analytically to real time at the end if needed. In this case one often finds that all $d+1$ dimensions are equivalent. In this discussion of the scalar field we shall assume that this is so.

[2]In the fermion problem we are going to study, the (imaginary) time and space directions are not equivalent. The correlation length in these discussions refers to the time direction.

By assumption we are going to be interested only in correlations of $\phi_<$. However, at present we are computing these objects by doing an integral over fast and slow modes. The first step in the RG program is to ask if there is an effective action or Boltzmann weight $e^{S'(\phi_<)}$ such that, when integrated over just the slow modes, it will reproduce all the slow correlation functions. We shall now see that the answer is affirmative.

Let the action be expressed as follows:

$$S(\phi_<,\phi_>)=S_0(\phi_<)+S_0(\phi_>)+S_I(\phi_<,\phi_>)\ , \tag{20}$$

where S_0 is a quadratic function of its arguments that separates into slow and fast pieces and S_I, called the interaction, is the part which mixes the two. Then

$$Z=\int \prod_{0\le k<\Lambda/s} d\phi(k) \times \prod_{\Lambda/s\le k<\Lambda} d\phi(k) e^{S_0(\phi_<)} e^{S_0(\phi_>)} e^{S_I(\phi_<,\phi_>)} \tag{21}$$

$$\equiv \int [d\phi_<]\int [d\phi_>] e^{S_0(\phi_<)} e^{S_0(\phi_>)} e^{S_I(\phi_<,\phi_>)} \tag{22}$$

$$= \int [d\phi_<] e^{S_0(\phi_<)} \int [d\phi_>] e^{S_0(\phi_>)} e^{S_I(\phi_<,\phi_>)} \tag{23}$$

$$\equiv \int [d\phi_<] e^{S'(\phi_<)} \tag{24}$$

which defines the *effective action* $S'(\phi_<)$. Let us manipulate its definition a little:

$$\begin{aligned} e^{S'(\phi_<)} &= e^{S_0(\phi_<)} \int [d\phi_>] e^{S_0(\phi_>)} e^{S_I(\phi_<,\phi_>)} \\ &= e^{S_0(\phi_<)} \frac{\int [d\phi_>] e^{S_0(\phi_>)} e^{S_I(\phi_<,\phi_>)}}{\int [d\phi_>] e^{S_0(\phi_>)}} \\ &\quad \times \underbrace{\int [d\phi_>] e^{S_0(\phi_>)}}_{Z_{0>}} \\ &= e^{S_0(\phi_<)} \langle e^{S_I(\phi_<,\phi_>)} \rangle_{0>} \end{aligned} \tag{25}$$

where $\langle\ \rangle_{0>}$ denotes averages with respect to the fast modes with action S_0 and where the corresponding partition function $Z_{0>}$ has been dropped in going to the last line, since it will merely add a constant to the effective action independent of $\phi_<$, which in turn will make no difference to any correlation function of slow modes.

Although $S'(\phi_<)$ provides a good description of the slow-mode physics, the RG transformation has two more steps besides the above *mode elimination*. These steps will now be motivated.

One aim of the RG is to see the various parameters in the interaction evolve or flow as the cutoff is reduced, i.e., to compute the β function. Suppose before mode elimination we had

$$S(\phi)=r\phi^2+u\phi^4+\cdots \tag{26}$$

and after,

$$S'(\phi_<)=r'\phi_<^2+u'\phi_<^4+\cdots\ . \tag{27}$$

(The action above is schematic. For example $u\phi^4$ could be the shorthand for

$$\int dk_1 dk_2 dk_3 dk_4 \delta(k_4+k_3+k_2+k_1) \times u(k_4,k_3,k_2,k_1)\phi(k_4)\phi(k_3)\phi(k_2)\phi(k_1)\ , \tag{28}$$

where $u(k_4,\ldots,k_1)$ is a *coupling function* and not just a coupling constant.) In any event, we are trying to compare r to r', u to u', and so on. The problem with doing that is that we are comparing apples to oranges. The old and new theories are defined on two different kinematical regions. For example, the coupling $u(\Lambda,\Lambda,\Lambda,\Lambda)$ has no counterpart in the effective theory, which has all its momenta below Λ/s. (In field theory, where the old and new cutoff are both sent to infinity, this point does not come up.) To remedy this defect, we shall define new momenta after mode elimination,

$$k'=sk\ , \tag{29}$$

which run over the same range as k did before elimination.

There is just one more problem. Consider two actions:

$$S(\phi)=r\phi^2+u\phi^4\ , \tag{30}$$

$$S'(\phi)=4r\phi^2+16u\phi^4 \tag{31}$$

which seem different. They are, however, physically equivalent because we can simply define $2\phi=\phi'$ in the second action (and ignore the Jacobian in the functional integral since it is a ϕ-independent constant) and reduce it to the first action. In other words, certain changes in parameters are not of physical importance, since they can be absorbed by field rescaling. To weed these out, we shall follow mode elimination and momentum rescaling by a field rescaling, defining new fields,

$$\phi'(k')=\zeta^{-1}\phi_<(k'/s)\ , \tag{32}$$

and choose ζ such that a certain coupling in the quadratic part of the action has a fixed coefficient. The final action S' will then be expressed in terms of this new field. Thus the three stages in the RG transformation are as follows: Eliminate fast modes, i.e., reduce the cutoff from Λ to Λ/s. Introduce rescaled momenta $k'=sk$, which now go all the way to Λ. Introduce rescaled fields $\phi'(k')=\zeta^{-1}\phi_<(k'/s)$ and express the effective action in terms of them. This action should have the same coefficient for a certain quadratic term.

B. Fixed points

With this definition of the RG transformation, we have a mapping from Hamiltonians or actions defined in a certain phase space to actions in the same space. Thus, if we represent the initial action as a point in a coupling constant space, this point will flow under the RG transformation to another point in the same space. This

definition of the RG opens up a possibility that did not exist without all three steps: *a fixed point* S^* of the group action, that is to say, the action function, which reproduces itself after the three-step RG transformation. Geometrically this means that the point S^* does not move or flow under the RG action. If the system had a correlation length ξ before the RG (in the old units), in the new units (in which momenta get boosted by a factor s) it would decrease to ξ/s. On the other hand, at the fixed point, it must remain the same under the RG. This means that the correlation length at a fixed point must have been either zero or infinite. We shall be dealing with the latter case here. Fixed points will dominate our analysis.

In summary, we see that in the modern viewpoint, the cutoff is not to be viewed as an artifact to be sent to infinity but as the dividing line between the modes we are interested in and the modes we are not interested in. The preceding discussion explains *how* we may change the cutoff and the couplings without affecting the slow-mode physics even in a problem where there were no ultraviolet infinities. Let us now understand *why* we would want to do such a thing.

Consider the remarkable phenomenon of universality. How can systems with different microscopic Hamiltonians have the same decay exponent x in their critical two-point functions? The RG explains this as follows (Kadanoff, 1965, 1977; Wilson, 1971, 1975; Fisher, 1974, 1983; Wilson and Kogut, 1974). Let S_A and S_B be two critical Hamiltonians defined in the full k space. Each is described by a set of coupling constants. Let us represent each as a point in a space $\mathcal{S}$ (in the notation of Wilson and Kogut, 1974, Section 12) of Hamiltonians where along each axis we measure one coupling constant. The fact that $S_A \neq S_B$ implies that they are given by distinct points in coupling-constant space and that there are many observables that differ in the two cases. Consider, however, extreme long-distance physics, in particular, the long-distance decay of two-point functions. To calculate these we need just the slow modes. To this end let us trade each Hamiltonian for its equivalent one after renormalization down to a very small cutoff. *What we shall find is that they both asymptotically approach the same fixed-point Hamiltonian* S^* *where the flow stops.* This explains why they share the same long-distance physics and in particular the exponent x. Although the coupling-constant space is infinite dimensional, let us consider a toy model in which it is three dimensional. Let all critical Hamiltonians (in particular, S_A and S_B) lie in the $x-y$ plane. Under the RG they flow to S^*, which lies at, say, the point (1,1,0). Let us shift the origin of coordinates to the fixed point. Any deviation $S-S^*$ that lies in the critical plane is termed *irrelevant* in the RG terminology, since it renormalizes to zero and hence makes no difference to long-distance physics. The fixed point, being a special case of a critical point, will of course have a power-law decay of correlations.

What we see is that if to this fixed point an irrelevant perturbation is added, the perturbed system will also have the same power-law decay. If the functional integral stands for some quantum system written as a path integral, this means that a gapless system will remain gapless if an irrelevant perturbation is added. This idea will be invoked later in this article. By contrast, any deviation off the critical $(x-y)$ plane is called *relevant* and will get amplified by the RG transformation. The long-distance behavior of correlations in this problem is unclear; it is controlled by the ultimate destination of this flow and typically (but not always) corresponds to exponential decay. In the general problem there can also be *marginal* perturbations, which neither grow nor decay under the RG transformation. They play a major role in the nonrelativistic electron problem to which we now turn our attention. A truly marginal perturbation does not cause a gap.

Table I summarizes some of the above concepts.

TABLE I. Definitions of terms frequently used in connection with the RG.

Symbol	Meaning
RG	Renormalization group.
Z	Classical partition function or Feynman's path integral for quantum problem.
Λ	The cutoff, the maximum allowed value of momentum k.
s and t	The parameter in the RG: $\Lambda=\Lambda_0/s=\Lambda_0 e^{-t}$, where Λ_0 is fixed.
Slow modes $\phi_<$	Modes to be retained.
Fast modes $\phi_>$	Modes to be integrated out.
β function	The rate of change of couplings with t, the logarithm of the cutoff.
S	The action or Hamiltonian. The Boltzmann weight is e^S.
$\mathcal{S}$	The space of all Hamiltonians. Each axis is used to measure one parameter.
S_A or S_B	Any two actions or Hamiltonians describing two different systems. Points in $\mathcal{S}$.
Critical system	A system tuned to be at a phase transition. Has power-law correlations.
Critical exponents	Exponents for the above power laws; universal.
Critical surface	The locus of all actions or Hamiltonians that describe critical systems.
S^*	The fixed point of the RG transformation in the space of Hamiltonians.
Relevant variable	Any deviation from S^* which gets amplified under the RG action.
Irrelevant variable	Any deviation from S^* which gets renormalized to zero.
Marginal variable	Any deviation from S^* which remains fixed under the action of the RG.

C. Central problem of this review

The preceding discussions have set the stage for introducing our main topic. Consider a system of noninteracting fermions[3] at zero temperature ($T=0$), either in the continuum with a dispersion relation

$$E=K^2/2m \tag{33}$$

or on a lattice with some energy function $E(\mathbf{K})$ defined within the Brillouin zone.[4] In all cases, one-particle states with $E\leq\mu$, where μ is the *chemical potential,* are filled in the ground state of the many-body system. The filled states are bounded by the *Fermi surface.* In the continuum in $d=2$ or $d=3$ the Fermi surface is a circle or sphere, respectively, of radius

$$K_F=\sqrt{2m\mu}\ . \tag{34}$$

This ground state has gapless excitations corresponding to the promotion of fermions from just below the Fermi surface to just above it. The central questions we ask in this paper are the following: If some perturbation is added to the free theory, will the system develop a gap at once or will it remain gapless? If it remains gapless, what is the natural way to describe the low-energy physics of the system, in particular its response to "soft probes," probes of low-frequency ω and momentum $\mathbf{Q}$? The answers to these questions are clearly dictated by the modes near the Fermi surface, at least for the case of weak perturbations. For example, in any kind of perturbation theory of the ground state, these modes will come with the lowest-energy denominators. We shall therefore focus on modes within a bandwidth Λ of the Fermi surface (the slow modes of this problem) and get rid of all the modes outside this cutoff (the fast modes). In the case of fermions in free space we define a lower-case momentum

$$k=|\mathbf{K}|-K_F \tag{35}$$

and work with modes obeying

$$|k|\leq\Lambda\ . \tag{36}$$

In the case of electrons on a lattice, the wave vector $|\mathbf{K}|$ is no longer a measure of energy, and we must keep those modes whose *energy* lies within some cutoff. This complication will be discussed in the sections devoted to lattice problems. For the present let us focus on electrons in free space and imagine an annulus or shell (in two or three dimensions, respectively) of thickness 2Λ with mean radius K_F, within which reside the slow modes of this problem.

Let us now turn to the elimination of the fast modes outside the cutoff. This may be done within the operator formalism by the use of projection operators to define an effective Hamiltonian restricted to the subspace of slow modes. This effective quantum Hamiltonian depends on the cutoff in such a way as to produce cutoff-independent results for the surviving slow modes, and the fixed point, if any, is unaffected by this transformation. This is the approach used by Anderson and Yuval (1970), Nozières (1974), Wilson (1975), and Kirshna-Murthy *et al.* (1980) in their treatment of the Kondo problem. However, this problem, which is a paradigm for how the RG is to be used in quantum problems in many-body physics, is essentially one dimensional.[5] By contrast the problems we deal with here are truly two and three dimensional, and the application of the RG to these has a short history. Although Anderson had suggested this possibility in his book (Anderson, 1984), no detailed analysis was carried out for some time. Benfatto and Gallavotti (1990) and Feldman *et al.* (Feldman and Trubowitz, 1990, 1991; Feldman *et al.*, 1992, 1993) then combined the RG with rigorous bounds to study (to all orders in perturbations) the stability of gapless Fermi systems to perturbations. Shankar (1991) developed the method to be described here, which is less rigorous, more intuitive, and covers other instabilities like charge or spin-density waves and rotationally-noninvariant systems and which may be easier to use for people with a background in critical phenomena or modern field theory. More recently Polchinski (1992) employed a very similar approach to the nonrelativistic fermion problem to better understand the concept of effective field theories in particle physics. Weinberg (1993) recently derived the effective low-energy action and RG flow equations for superconductors with Fermi surfaces that obeyed time-reversal symmetry and nothing else. All these approaches are fundamentally different in spirit from the method used by Hertz (1976), who completely integrated the fermions in favor of some bosonic variables. In particular, he integrated the modes at the Fermi surface. This is analogous to integrating the $k=0$ modes in critical phenomena. The effective theory for the bosons then has singular parameters. Hertz found a way to analyze phase transitions that can be described by the bosonic variables. For a recent analysis of Hertz's approach see Millis (1993) and Sachdev (1993).

The approach described in this paper is as follows. To heighten the analogy with critical phenomena one first

[3]We use the term fermion instead of simply electron to accommodate *spinless fermions,* which do not exist in nature but simplify the analysis by obviating the need for spin indices. While the spin of the electron is certainly not ignorable when comparing theory to experiment, it will be seen that it really is an inessential complication in the RG program to be described here and may be incorporated readily.

[4]We use upper-case letters to denote momenta measured from the origin in contrast to the preceding discussion, where lower-case symbols were used. This is a deliberate departure from convention and reflects the different physics that emerges here.

[5]Here one deals with a conduction band of electrons interacting with a single fixed impurity. By using spherical waves (instead of plane waves) centered on the impurity and keeping just the s wave, one reduces it to a quasi-one-dimensional problem in the radial coordinate.

shifts from the operator approach to a path-integral approach and writes down the path integral. First the noninteracting problem is considered. Since it is gapless one expects it to be the fixed point of a RG transformation in which the cutoff is reduced. Such a transformation is found. With respect to this transformation, perturbations are classified as relevant, irrelevant, or marginal. In the last two cases the system will remain gapless, while in the first case one can only make statements if one assumes that the behavior seen for small perturbations persists at strong coupling also. In all cases considered, this corresponds to the opening of a gap in the spectrum. While all this sounds like critical phenomena (and is meant to), there are crucial differences, which can be traced back to the nature of the phase space for slow modes. In critical phenomena all the action at long wavelengths is focused on a tiny *ball centered at the origin of k space,* and the fixed point is characterized by a few couplings.[6]

The same is true for continuum field theories like quantum electrodynamics or quantum chromodynamics: both fermion and boson momenta are restricted to a sphere of radius Λ centered at the origin. In the problem at hand, we renormalize not towards a single point, the origin, but towards a surface, the Fermi surface (which may itself change under renormalization in the nonspherical case). In contrast to critical phenomena, where all momenta and momentum transfers are small (bounded by the cutoff), here only $k = |\mathbf{K}| - K_F$ is small, and large momentum transfers of the order of K_F are possible within the slow modes. *Renormalization only reduces the dimension normal to the Fermi surface; the tangential part survives.* As for the fixed point, it is characterized by a surface and *coupling functions* defined on it.[7] Notice that $d = 1$ is special: here the Fermi surface is a set of two disjoint points. Apart from this doubling (which converts nonrelativistic fermions into Dirac fermions), we have the same situation as in a continuum field theory in one space dimension, and there are once again just a few coupling constants. This is why there has been a lot of activity and a lot of success (Solyom, 1979; Bourbonnais and Caron, 1991) in applying the RG to one-dimensional fermion problems in condensed matter and a lot of resistance to going to higher dimensions.

D. Outline

We now turn to the details. In Sec. II the reader is given a very brief review of how the RG works for a scalar field theory in four dimensions. This will serve to remind readers familiar with the subject of the highlights that we shall recall frequently in our progress by analogy. As for the newcomers, it will give them the minimum required to follow this article. References for more details will be given.

Section III explains how a path integral can be written for fermions and how one is to extract correlation functions from it. This will require the introduction of Grassmann variables. Readers not used to these should kill two birds with one stone by using the pedagogical review provided here to learn this tool, which is often used in condensed-matter theory.

In Sec. IV we study the problem of spinless fermions in one dimension at half-filling: with one particle per every other site on the average. This section serves as a warm up for the RG program, since, as explained above, it resembles the run-of-the-mill field theory in one dimension. It also shows the power of the RG: whereas mean-field theory (a self-consistent approximation to be detailed later) predicts a gap for the smallest repulsion, and superconductivity for the smallest attraction, the exact solution tells us that the system remains gapless for a finite range of coupling of either sign. It will be seen that the RG gives results in agreement with the exact solution.

Sections V and VI deal with circular and spherical Fermi surfaces. To lowest order in a perturbative expansion (Sec. V) one finds that there exists a fixed point described by two marginal coupling functions F and V, which depend on the angles on the circle or sphere as the case may be. To the next order (Sec. VI) one finds that F is still marginal, while each coefficient in the angular momentum expansion of V grows to produce the superconducting instability if attractive, and renormalizes to downwards if repulsive, a result originally discovered by Morel and Anderson (1962). No new surprises come at higher orders. This is explained in the next section. The fixed-point theory, which exists in the absence of V, is what is known as Landau's Fermi-liquid theory. The Kohn-Luttinger effect, which destroys the Fermi liquid at low temperatures, is derived in the RG language.

Section VII provides a new way of understanding why it is possible to solve the fixed-point theory characterized by the interaction F even though F is not necessarily small. This is tied to the fact that certain theories with a large number of fields can be described by an expansion in $1/N$, N being the number of components. (In other words, the coupling need not be small as long as $1/N$ is.) It is shown that the Fermi system with cutoff Λ has a

[6] I thank Pierre Hohenberg for pointing out to me an exception—a problem not involving a Fermi surface, which has nonetheless a similar phase space after any amount of renormalization: the condensation of a liquid into a nonuniform state, studied by Brazovskii (1975). See Swift and Hohenberg (1977) for the study of fluctuations on an equivalent model. It is an open question whether the methods developed here can be applied to Brazovskii's problem.

[7] All this can be stated in another way. In field theories or in critical phenomena one also runs into coupling functions. But these are functions just of k. When Taylor expanded in k, only a few terms are marginal or relevant. In the present problem, the coupling functions depend on k as well as the coordinates of the limiting Fermi surface. The latter never get eliminated, and all terms in the Taylor series for the latter will be important. This point will be discussed further as we go along.

$1/N$ expansion with $N=K_F/\Lambda$. Thus a given problem in the full momentum space can initially be renormalized to a small Λ theory (without running into any singularities) and then, when N is large enough, solved in the $1/N$ approximation.

Section VIII has a discussion of Landau's Fermi-liquid theory. Only some aspects of this extensive field are brought up.

In Sec. IX we consider noncircular Fermi surfaces with no special features other than time-reversal invariance: if $\mathbf{K}$ lies on it so does $-\mathbf{K}$. It is found that one must stop using $|K|$ as a measure of energy and use actual equal-energy contours to define the fast and slow modes. The net result is exactly as in the rotationally invariant case except for the fact that F and V now depend on more variables due to the lack of rotational invariance.

Section X deals with the very interesting case of *nested* Fermi surfaces in $d=2$: surfaces such that, if $\mathbf{K}$ lies on them, so does $\mathbf{K}+\mathbf{Q}_N$, where $\mathbf{Q}_N$ is a fixed *nesting* momentum. We choose to illustrate the ideas with spinless fermions on a rectangular lattice, in which case the nesting vector $\mathbf{Q}_N$ has components (π,π). (Readers unfamiliar with nesting may wish to peek at Fig. 17 in Sec. X for an example.) What we find is that to lowest order a third coupling function W insinuates itself at the fixed point. At next order it begins to flow. One can show that there are definitely some relevant directions if this force is repulsive, and these tend to produce charge-density waves: the ground state has a nonuniform charge density which oscillates with momentum $\mathbf{Q}_N$.

In Sec. XI we use the methods developed here to look for non-Fermi liquids in two dimensions. Regrettably the results are negative for the case of weakly coupled problems with a circular Fermi surface.

Section XII contains the summary and outlook. Many of the remarks made in this preview will be repeated there, and the reader will have a clearer picture of their significance. Two Appendices deal with special topics. Coulomb screening and the Kohn-Luttinger effect as they appear within the RG framework.

II. AN EXAMPLE OF THE RG FROM $d=4$

The problem chosen to illustrate the RG at work involves a complex scalar field in $d=4$. The functional integral can be viewed either as describing the quantum field theory of a charged scalar field in three space dimensions or as describing the classical statistical mechanics of a system with one complex field or two real fields at each point on a lattice. The Ising model which is described by a real field is not chosen here, since the fermion problem we shall study later involves charged Fermi fields. Readers new to the problem should be aware that this section has the very limited objective of making the rest of the paper comprehensible to them. For a deeper introduction to critical phenomena, the reader is directed to any of the excellent reviews (Fisher, 1974; Wilson and Kogut, 1974; Kadanoff, 1977) or books (Ma, 1976; Itzykson and Drouffe, 1989; Plischke and Bergersen, 1989; Zinn-Justin, 1989; Le Bellac, 1991; and Goldenfeld, 1992). Readers familiar with the subject are still urged to skim through this section to get acquainted with the notation and to refresh their memory, since our approach to the interacting Fermi problem will rely heavily on analogy to this problem, for which the RG approach has been spectacularly successful.

The partition function for this problem is

$$Z=\int_{|k|<\Lambda}[d\phi d\phi^*]e^{S(\phi,\phi^*)}\,, \tag{37}$$

where

$$[d\phi d\phi^*]=\prod_{|k|<\Lambda}\frac{d\,\mathrm{Re}\phi(\mathbf{k})d\,\mathrm{Im}\phi(\mathbf{k})}{\pi}\,, \tag{38}$$

$$S(\phi,\phi^*)=\int_{|k|<\Lambda}\phi^*(\mathbf{k})J(k)\phi(\mathbf{k})\frac{d^4k}{(2\pi)^4} \tag{39}$$

and

$$J(k)=2[(\cos k_x-1)+(\cos k_y-1)+(\cos k_z-1)+(\cos k_t-1)]\,. \tag{40}$$

This action is obtained by Fourier transformation of the following nearest-neighbor interaction in coordinate space:

$$S(\phi,\phi^*)=-\frac{1}{2}\sum_{\mathbf{n},\mathbf{i}}|\phi(\mathbf{n})-\phi(\mathbf{n}+\mathbf{i})|^2\,, \tag{41}$$

where $\mathbf{n}$ is the vector with integer coordinates used to label the sides of the hypercubic lattice and $\mathbf{i}$ is any of the eight unit vectors in the direction of increasing or decreasing coordinates. Notice that this action favors the alignment of neighboring fields, i.e., is ferromagnetic.

Since we are interested in small-k physics, let us hereafter approximate $J(k)$ by its leading term in the Taylor series and write

$$S_0(\phi,\phi^*)=-\int_{|k|<\Lambda}\phi^*(\mathbf{k})k^2\phi(\mathbf{k})\frac{d^4k}{(2\pi)^4}\,, \tag{42}$$

$$Z=\int_{|k|<\Lambda}[d\phi\,d\phi^*]e^{-\int_{|k|<\Lambda}\phi^*(\mathbf{k})k^2\phi(\mathbf{k})d^4k/(2\pi)^4}\,, \tag{43}$$

$$[d\phi\,d\phi^*]=\prod_{|k|\le\Lambda}\frac{d\,\mathrm{Re}\phi(\mathbf{k})d\,\mathrm{Im}\phi(\mathbf{k})}{\pi}=\prod_{|k|\le\Lambda}\frac{d\phi^*(\mathbf{k})d\phi(\mathbf{k})}{2\pi i}\,. \tag{44}$$

This is called the *Gaussian model.* The corresponding functional integrals are the product of ordinary Gaussian integrals, one for each $\mathbf{k}$. This makes it possible to express all the correlation functions in terms of averages involving a single Gaussian integral. *The only averages that do not vanish are products of an even number of variables, wherein each $\phi(\mathbf{k})$ is accompanied by its complex conjugate.* This is because the action and measure are invariant under

$$\phi(\mathbf{k})\to\phi(\mathbf{k})e^{i\theta},\quad \phi^*(\mathbf{k})\to\phi^*(\mathbf{k})e^{-i\theta}, \tag{45}$$

where θ can be different for different $\mathbf{k}$'s. Thus the only integral we shall ever need follows from the simple problem involving just a pair of complex-conjugate variables z and z^*:

$$\langle zz^*\rangle=\frac{\int_{-\infty}^{\infty}(dz\,dz^*/2\pi i)zz^*e^{-azz^*}}{\int_{-\infty}^{\infty}(dz\,dz^*/2\pi i)e^{-azz^*}}=\frac{1}{a}. \tag{46}$$

The other two bilinears have zero average:

$$\langle zz\rangle=\langle z^*z^*\rangle=0 \tag{47}$$

because the action and measure are invariant under

$$z\to ze^{i\theta},\quad z^*\to z^*e^{-i\theta}, \tag{48}$$

while the bilinears are not. The reader wishing to verify the above results is asked to switch to x and y, the real and imaginary parts of the integration variables, and to use

$$\frac{dz\,dz^*}{2\pi i}\to\frac{dx\,dy}{\pi}. \tag{49}$$

If there are two sets of variables we have

$$\langle z_i^*z_j\rangle=\frac{\int_{-\infty}^{\infty}\frac{dz_1dz_1^*}{2\pi i}\frac{dz_2dz_2^*}{2\pi i}z_i^*z_je^{-a_1z_1z_1^*-a_2z_2z_2^*}}{\int_{-\infty}^{\infty}\frac{dz_1dz_1^*}{2\pi i}\frac{dz_2dz_2^*}{2\pi i}e^{-a_1z_1z_1^*-a_2z_2z_2^*}}$$

$$=\frac{\delta_{ij}}{a_i}\equiv\langle\bar{i}j\rangle. \tag{50}$$

As for the four-point function, the reader may verify that

$$\langle z_i^*z_jz_k^*z_l\rangle=\langle\bar{i}j\rangle\langle\bar{k}l\rangle+\langle\bar{i}l\rangle\langle\bar{k}j\rangle. \tag{51}$$

This result makes sense. It demands that, for the answer to be nonzero, the fields come in complex-conjugate pairs. Since this can happen in two ways, the result is a sum of two terms. The generalization to more variables and longer strings is obvious.

In view of the above, the reader will not be surprised that the *two-point function* in our Gaussian model is

$$\langle\phi^*(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle=\frac{(2\pi)^4\delta^4(\mathbf{k}_1-\mathbf{k}_2)}{k_1^2} \tag{52}$$

$$\equiv(2\pi)^4\delta^4(\mathbf{k}_1-\mathbf{k}_2)G(k_1) \tag{53}$$

$$\equiv\langle\bar{2}1\rangle, \tag{54}$$

and likewise

$$\langle\phi^*(\mathbf{k}_4)\phi^*(\mathbf{k}_3)\phi(\mathbf{k}_2)\phi(\mathbf{k}_1)\rangle=\langle\bar{4}2\rangle\langle\bar{3}1\rangle+\langle\bar{4}1\rangle\langle\bar{3}2\rangle. \tag{55}$$

This is a case of *Wick's theorem* for bosons. For the case of $2n$ fields, the answer is a sum over all possible pairings, each term in the sum being a product of n two-point functions. The result follows from the preceding discussion, upon making the change from Kronecker deltas to Dirac delta functions in Eqs. (51) to take into account the fact that the action in the Gaussian model is an integral (over $\mathbf{k}$) rather than a sum over variable labels.

Note that G has power-law behavior in momentum space $(1/k^2)$ and hence will do so in coordinate space $(1/r^2)$. Thus the action of the Gaussian model is critical and must flow to a fixed point under the action of the RG. We shall now see that it is itself a fixed point.

A. RG for free field

In the first stage of the RG transformation, we integrate out $\phi_>$. Since $S_I=0$ here, we see from Eq. (25) that

$$S'(\phi_<)=-\int_{|k|<\Lambda/s}\phi_<^*(\mathbf{k})k^2\phi_<(\mathbf{k})\frac{d^4k}{(2\pi)^4}. \tag{56}$$

We now carry out the last two steps by rewriting the action in terms of

$$\phi'(\mathbf{k}')=\zeta^{-1}\phi_<(\mathbf{k}'/s) \tag{57}$$

and obtain

$$S'(\phi'_<)=-s^{-6}\int_{|k'|<\Lambda}\phi_<^*(\mathbf{k}'/s)k'^2\phi_<(\mathbf{k}'/s)\frac{d^4k'}{(2\pi)^4} \tag{58}$$

$$=-\frac{\zeta^2}{s^6}\int_{|k'|<\Lambda}\phi'^*(\mathbf{k}')k'^2\phi'(\mathbf{k}')\frac{d^4k'}{(2\pi)^4} \tag{59}$$

$$=S'(\phi'). \tag{60}$$

If we now make the choice

$$\zeta=s^3 \tag{61}$$

we find that the Gaussian action is the fixed point:

$$S'(\phi')=S(\phi)=S^*. \tag{62}$$

Having found the fixed point, we next classify its perturbations as relevant, irrelevant, or marginal. We shall consider only perturbations involving an even number of fields. Let us start with the quadratic case,

$$\delta S=-\int_{|k|<\Lambda}\phi^*(\mathbf{k})r(k)\phi(\mathbf{k})\frac{d^4k}{(2\pi)^4}, \tag{63}$$

where the *coupling function* r is assumed to have a Taylor expansion

$$r(k)=r_0+r_2k^2+\cdots \tag{64}$$

which reflects the short-range nature of the perturbation in coordinate space. One often writes

$$r_0=m_0^2 \tag{65}$$

and refers to m_0^2 as the *mass term,* since in the quantum field theory interpretation of the functional integral, adding this term to the Gaussian model yields the quantum field theory of a particle of mass m_0.

Since this perturbation does not mix slow and fast modes, all we have to do is replace ϕ by $\phi_<$ everywhere and reexpress the result in terms of new momenta and fields. This gives

$$\delta S'(\phi'(\mathbf{k}'))=-s^2\int_{|k|'<\Lambda}\phi'^*(\mathbf{k}')r(k'/s)\phi'(\mathbf{k}')\frac{d^4k'}{(2\pi)^4} \tag{66}$$

$$\equiv-\int_{|k|'<\Lambda}\phi'^*(\mathbf{k}')r'(k')\phi'(\mathbf{k}')\frac{d^4k'}{(2\pi)^4}\ , \tag{67}$$

where in the last equation we have invoked the definition of the renormalized coupling $r'(k')$. By comparison, we find

$$r'(k')=s^2r(k'/s)\ , \tag{68}$$

which implies that the Taylor coefficients obey

$$r_0'=s^2r_0\ , \tag{69}$$

$$r_2'=r_2\ , \tag{70}$$

$$r_4'=s^{-2}\tau_4\ , \tag{71}$$

and so on. Thus we find that r_0 is relevant, r_2 is marginal, and the rest are irrelevant. This is a concrete example of how in the low-energy physics the coupling function $r(k)$ reduces to a few coupling constants. (In fact r_2 makes no difference, since it can be absorbed by field rescaling.) In quantum field theory, where we send the cutoff to infinity, all momenta are small compared to the cutoff and the theory is defined by a few coupling constants. We shall see that the same thing will happen for the quartic interaction: a coupling function of four different momenta will reduce to a single coupling constant. We may understand all this as follows. In the original Brillouin zone, of size $1/a$, all these functions are nontrivial and we need them in their entirety. As we eliminate modes, we need their behavior in a smaller and smaller ball near the origin; see Eq. (68). Not surprisingly, the function is well described by a few terms in the Taylor series. This is the picture in fixed or "laboratory units." In the RG one uses sliding units that constantly change to keep the cutoff (ball size) fixed at Λ, and the same phenomenon appears as the rapid shrinkage of higher coefficients in the Taylor series. (Of course, as we renormalize, we are not just rewriting the original coupling function in new units; the function itself changes due to eliminated modes. But it is expected nonetheless to be smooth in k. This is one of the points emphasized in the modern RG theory: elimination of modes does not introduce new singularities into the couplings. As we shall see, this is because the effect of mode elimination may be expressed in terms of integrals which are convergent in the infrared and ultraviolet.)

B. Quartic perturbation

Let us now consider the quartic perturbation

$$\delta S=-\frac{1}{2!2!}\int_{|k|<\Lambda}\phi^*(\mathbf{k}_4)\phi^*(\mathbf{k}_3)\phi(\mathbf{k}_2)\phi(\mathbf{k}_1)u(\mathbf{k}_4\mathbf{k}_3\mathbf{k}_2\mathbf{k}_1)(2\pi)^4\delta^4(\mathbf{k}_4+\mathbf{k}_3-\mathbf{k}_2-\mathbf{k}_1)\prod_{i=1}^{4}\frac{d^4k_i}{(2\pi)^4}$$
$$\equiv-\int_\Lambda\phi^*(4)\phi^*(3)\phi(2)\phi(1)u(4321)\ , \tag{72}$$

where the coupling function obeys the symmetry condition

$$u(4321)=u(3421)=u(4312)\ . \tag{73}$$

In other words, the coupling function is invariant under the exchange of the first two or last two arguments. Even if we started with a function that did not have this symmetry, the invariance of the measure and the rest of the integrand under this symmetry would automatically project out the symmetric part and annihilate the antisymmetric part. The factorials up front are conventional and are put there to prevent similar factors from arising in subsequent calculations.

The renormalization of the quartic interaction is complicated by the fact that, unlike the quartic perturbations, it mixes up the slow and fast modes. Thus we have to use the formula

$$e^{S'(\phi_<)}=e^{S_0(\phi_<)}\langle e^{\delta S(\phi_<,\phi_>)}\rangle_{0>} \tag{74}$$

$$\equiv e^{S_0+\delta S'} \tag{75}$$

Next we invoke the *cumulant expansion,* which relates the mean of the exponential to the exponential of means:

$$\langle e^{\Omega}\rangle=e^{[\langle\Omega\rangle+(\langle\Omega^2\rangle-\langle\Omega\rangle^2)/2+\cdots]}\ . \tag{76}$$

The reader may wish to verify the correctness of this expansion to the order shown. Using this expansion we find

$$\delta S'=\langle\delta S\rangle+\tfrac{1}{2}(\langle\delta S^2\rangle-\langle\delta S\rangle^2)+\cdots \tag{77}$$

Since δS is linear in u, this is a weak-coupling expansion. It is now clear what has to be done. Each term in the series contains some monomials in fast and slow modes. The former have to be averaged with respect to the Boltzmann weight $S_0(\phi_>)$ by the use of Wick's theorem. The result of each integration will be to give a monomial in the slow modes. When reexpressed in terms of the rescaled fields and momenta, each will renormalize the corresponding coupling. In principle the reader has been given information to carry out this process. There is,

however, no need to reinvent the wheel. There is a procedure involving Feynman diagrams which automates this process. These rules will not be discussed here since they may be found, for example, in Secs. III–V of Wilson and Kogut (1974) or in any good field theory book (Itzykson and Zuber, 1980; Zinn-Justin, 1989). Instead we shall go over just the first term in the series in some detail and comment on some aspects of the second term. Readers familiar with Feynman diagrams should note that, while these diagrams have the same multiplicity and topology as the field theory diagrams, the momenta being integrated out are limited to the shell being eliminated, i.e., $\Lambda/s<k<\Lambda$.

The leading term has the form

$$\langle \delta S \rangle = -\frac{1}{2!2!}\left\langle \int_{|k|<\Lambda} (\phi_>+\phi_<)_4^*(\phi_>+\phi_<)_3^*(\phi_>+\phi_<)_2(\phi_>+\phi_<)_1 u(4321)\right\rangle_{0>} . \tag{78}$$

The sixteen possible monomials fall into four sets: 8 terms with an odd number of fast fields; 1 term with all fast modes; 1 term with all slow modes (called the *tree-level* term); 6 terms with two slow and two fast modes.

We have no interest in the first two sets, the first since it vanishes by symmetry and the second since it makes a constant contribution, independent of $\phi_<$, to the effective action. Consider next the third set with all slow modes, distinguished by the fact that it requires no integration (or averaging) over fast modes. This is called the *tree-level term* in field theory. The tree-level term is obtained from the original perturbation by simply setting $\phi=\phi_<$. Rewriting it in terms of new momenta and fields, we find it leads to the following quartic renormalized interaction:

$$\delta S'_{4,\text{tree}} = -\frac{1}{2!2!}\int_{|k|<\Lambda} \phi'^*(k'_4)\phi'^*(k'_3)\phi'(k'_2)\phi'(k'_1)u(k'_4/s,\ldots,k'_1/s)(2\pi)^4\delta^4(\mathbf{k}'_4+\mathbf{k}'_3-\mathbf{k}'_2-\mathbf{k}'_1)\prod_{i=1}^{4}\frac{d^4k'_i}{(2\pi)^4} . \tag{79}$$

The reader should note that the field rescaling factor s[12] has been exactly canceled by rewriting the delta function and integration measure in terms of new moments. (Note that the delta function scales oppositely to the momenta.)

It is evident that the renormalized four-point coupling is given by

$$u'(k'_4,\ldots,k'_1)=u(k'_4/s,\ldots,k'_1/s) . \tag{80}$$

Carrying out the Taylor expansion

$$u=u_0+O(k) , \tag{81}$$

we see that the constant term is marginal,

$$u'_0=u_0 , \tag{82}$$

and the rest are irrelevant. This is why the scalar field theory in four dimensions is described by a coupling constant and not a coupling function. Hereafter we shall replace the coupling function by the coupling constant. The effect will be irrelevant in the RG sense.

We now pass from the tree-level term to the six terms that have two slow and two fast modes in them. Of these, two with $\phi_>\phi_>$ or their conjugates are zero. The others clearly renormalize the quadratic interaction:

$$\delta S'_2(\phi_<) = -\frac{1}{2!2!}u_0\left\langle \int_{|k|<\Lambda}[\phi^*_>(4)\phi^*_<(3)+\phi^*_<(4)\phi^*_>(3)][\phi_>(2)\phi_<(1)+\phi_<(2)\phi_>(1)]\right\rangle_{0>} . \tag{83}$$

If we now evaluate the averages of the fast modes we shall find that all four terms give the same contribution (which takes care of the factorials in front), and we end up with

$$\delta S'_2(\phi_<) = -u_0\int_{|k|<\Lambda/s}\frac{d^4k}{(2\pi)^4}\phi^*_<(\mathbf{k})\phi_<(\mathbf{k})\int_{\Lambda/s}^{\Lambda}\frac{d^4k}{(2\pi)^4}\frac{1}{k^2} , \tag{84}$$

$$\delta S'_2(\phi'(\mathbf{k}')) = -u_0 s^2\int_{|k|<\Lambda}\frac{d^4k}{(2\pi)^4}\phi'^*(\mathbf{k}')\phi'(\mathbf{k}')\Lambda^2\frac{1}{2}\left[1-\frac{1}{s^2}\right]\frac{2\pi^2}{(2\pi)^4} , \tag{85}$$

where in the last step we have used the fact that the area of a unit sphere in four dimensions is $2\pi^2$.

Equation (85) gives us the change in r_0:

$$\delta r_0 = \frac{u_0\Lambda^2}{16\pi^2}(s^2-1) . \tag{86}$$

Let us agree to measure r_0 in units of the cutoff squared and drop the Λ^2 from now on.

Notice that the quartic coupling has renormalized the quadratic coupling. This is more the rule than the exception. The quadratic perturbations were special in that they did not generate new couplings. In view of this, we must really study the problem in which both r_0 and u_0 are present from the outset. This amounts to replacing the propagator $1/k^2$ by $1/(k^2+r_0)$ in Eq. (84). However, this only modifies the result to higher order in the expansion in r_0 and u_0. The flow to this order is

$$r'_0 = s^2\left[r_0+\frac{u_0}{16\pi^2}(1-1/s^2)\right] , \tag{87}$$

$$u_0' = u_0 \ . \tag{88}$$

If we take $s = 1+t$, with t infinitesimal, we find the differential equations

$$\frac{dr_0}{dt} = 2r_0 + \frac{u_0}{8\pi^2} \ , \tag{89}$$

$$\frac{du_0}{dt} = 0 \ . \tag{90}$$

This completes our analysis of the first term in the cumulant expansion. Let us see briefly how the above results follow in the diagrammatic approach. First we associate with each quartic perturbation δS a four-pronged X as in Fig. 1(a). The incoming arrows correspond to ϕ and the outgoing ones to ϕ^*. Each prong can stand for a $\phi_<$ or a $\phi_>$. Next we do the average over the fast modes. The prongs corresponding to the matching pairs that give a nonzero contribution are joined and correspond to the propagator. The diagrams in Fig. 1 tell us what happens. The first one corresponds to all slow modes, and there is nothing to average, i.e., no lines to join. Figure 1(b) corresponds to the eight terms with an odd number of fast lines. These average to zero. Figure 1(c) describes the case with two fast and two slow lines, with both sets coming in complex-conjugate pairs. The two fast lines are joined by the averaging, and the line joining them is the propagator; this corresponds to the renormalization of the quadratic term as per Eq. (84). This is called the *tadpole diagram.* Finally Fig. 1(d) describes the case in which all lines are fast and come in pairs. We now have two propagators. We did not consider this above, since it is a constant independent of $\phi_<$.

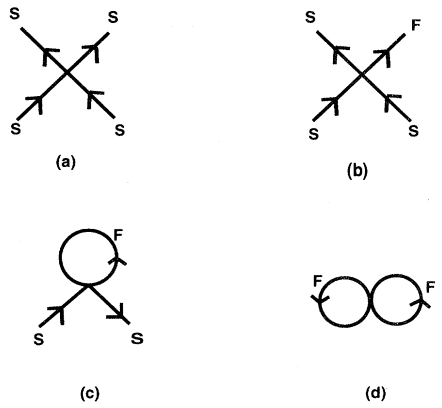

FIG. 1. Four types of diagrams that appear in the cumulant expansion to lowest order, corresponding to the $2^4 = 16$ choices for the legs on the vertex to be fast (F) or slow (S). (a) corresponds to all slow modes, denoted by S. This a *tree graph* with no integration over fast (F) modes. (b) typifies graphs that vanish, since they involve an odd number of fast lines. (c) is the *tadpole* graph. Here two fast lines have been joined to form a loop. This term with two external slow lines, will renormalize the quadratic term. (d) is a two-loop graph. It comes from a vertex with all fast lines upon joining them in pairs. Its contribution is a constant as far as the slow modes are concerned.

Notice that although all terms are of order u_0, they have very different topologies. The tree-level term has no loops or sum over fast modes. Figure 1(c) has one loop and Figure 1(d) has two loops. *Now the correct way to organize the cumulant expansion is by counting loops.* The reason is best seen in the language of quantum field theory, where the action has the prefactor $1/\hbar$ and the number of loops measures the powers of $\hbar$. In critical phenomena this fact becomes very clear when one works in $4-\varepsilon$ dimensions (Wilson and Fisher, 1972). One finds then that the loop expansion is an expansion in ε. The reader who wants to know more should consult the references given at the beginning of this section.

C. The one-loop graphs (ZS, ZS', BCS)

At zero loops, or tree level, the equations are

$$r_0' = s^2 r_0 \Longrightarrow \frac{dr}{dt} = 2r \ , \tag{91}$$

$$u_0' = u_0 \Longrightarrow \frac{du}{dt} = 0 \ . \tag{92}$$

Equations (89)–(90) are halfway between zero and one loop: they are good to one loop for r_0 and to tree level for u_0. To be consistent, we must evaluate the flow of u_0 to one loop also, which means going to second order in u_0 via the next term in the cumulant expansion, namely,

$$\tfrac{1}{2}[\langle(\delta S)^2\rangle - \langle \delta S)\rangle^2] \ .$$

Here we draw two crosses and do the usual pairing. All diagrams in which no line runs from one cross to the other, i.e., all disconnected diagrams, may be dropped since they get canceled by $-\langle(\delta S)\rangle^2$. Of the rest, the only graphs that affect u are shown in Fig. 2 and correspond to the following analytical expression:

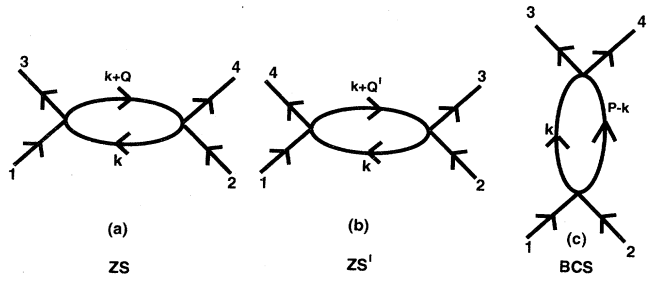

FIG. 2. The three one-loop graphs that renormalize the quartic coupling of the scalar field. The names ZS, ZS', and BCS are used to label the topology of the graphs and do not imply the corresponding phenomena (like BCS superconductivity). In (a) lines 1 and 3 meet at a vertex, in (b) 1 and 4 meet, while in the last lines 1 and 2 meet. All loop momenta lie in the shell being eliminated. The external momenta can be chosen to vanish if we want the renormalization of the marginal part of the quartic coupling. When this choice is made, both propagators have the same momentum. This is true for all three graphs.

$$u'(\mathbf{k}'_4,\ldots,\mathbf{k}'_1)=u_0-u_0^2\left[\underbrace{\int_{d\Lambda}\frac{d^4k}{(2\pi)^4}\frac{1}{k^2|\mathbf{k}-\mathbf{k}'_3/s+\mathbf{k}'_1/s|^2}}_{\text{ZS graph}}+\underbrace{\int_{d\Lambda}\frac{d^4k}{(2\pi)^4}\frac{1}{k^2|\mathbf{k}-\mathbf{k}'_4/s+\mathbf{k}'_1/s|^2}}_{\text{ZS' graph}}\right.$$
$$\left.+\frac{1}{2}\underbrace{\int_{d\Lambda}\frac{d^4k}{(2\pi)^4}\frac{1}{k^2|-\mathbf{k}+\mathbf{k}'_2/s+\mathbf{k}'_1/s|^2}}_{\text{BCS graph}}\right]. \tag{93}$$

Several remarks are in order. First note that, even though we started with a constant $u=u_0$, the renormalized coupling has acquired momentum dependence. If we expand the renormalized coupling in a Taylor series, keeping just the lowest term, we shall get the renormalized u_0. This is what we shall do, and ignore the irrelevant higher terms in the series. This in turn means that we can set all external momenta to zero. Before so doing, let us look at the three one-loop diagrams. Since we need to refer to them individually many times, we need a system of nomenclature. The one used here is by no means standard. Consider the first diagram, labeled ZS, which stands for "zero sound." In this diagram lines labeled 1 and 3 meet at a vertex. In Fermi-liquid theory a graph with the *same topology* occurs and is very important when $\mathbf{Q}=-\mathbf{k}_3+\mathbf{k}_1$ is small. The *physics* of the present problem could not be more different: the lines here stand for bosons and, unlike in Fermi-liquid theory, the internal loop momenta are restricted to lie at the cutoff rather than take all values within the cutoff. In the second ZS' diagram, lines 1 and 4 meet at a vertex. Usually when $\mathbf{Q}$ is small, $\mathbf{Q}'=-\mathbf{k}_4+\mathbf{k}_1$ is large. This diagram is not very important in Fermi-liquid theory and does not have a name. However, in problems with nesting, this diagram can be important if $\mathbf{Q}'$ is the nesting momentum. The BCS diagram with lines 1 and 2 meeting at a vertex has a topology like one that will appear later in our description of the superconducting instability. The reader is once again cautioned that the names of these diagrams are based solely on the topology and do not generally imply the corresponding physics.

Readers familiar with Feynman diagrams could have easily written them down. They must, however, pay attention to the symbol $\int_{d\Lambda}$, which reminds us that all internal propagator momenta (corresponding to integrated fast modes) are summed only over the modes being eliminated, which we take to be a shell of thickness $d\Lambda$ at the cutoff.

Readers new to the subject are strongly urged to work out the combinatorics and derive this result. They will then see why the factorials were included in the definition of the perturbation and why an extra factor of $\frac{1}{2}$ appears in the BCS diagrams.

All readers should note that the one-loop correction has a minus sign in front of it, reflecting the decrease of the interaction strength as we go to the infrared modes. (Although the one-loop graphs have a positive value, they reduce u_0, since the latter is defined to occur in the action with a negative sign; see Eq. (79).

Let us now set all external momenta to zero, since we are interested in just u_0. We are now assured that if the loop momentum $\mathbf{k}$ lies in $d\Lambda$, so does the other momentum which either equals $\mathbf{k}$ in the ZS and ZS' diagrams, or equals $-\mathbf{k}$ in the BCS case. All the integrals are now equal, and we get

$$u'_0=u_0-\frac{5u_0^2}{2}\int_{d\Lambda}\frac{k^3dk\,d\Omega}{(2\pi)^4k^4}, \tag{94}$$

$$\frac{du_0}{dt}=-\frac{5u_0^2}{16\pi^2}, \tag{95}$$

where in the last step we have recalled

$$\frac{|d\Lambda|}{\Lambda}=dt \tag{96}$$

and the area of a unit sphere in four dimensions ($2\pi^2$).

To one-loop accuracy we have the following flow:

$$\frac{dr_0}{dt}=2r_0+au_0, \tag{97}$$

$$\frac{du_0}{dt}=-bu_0^2, \tag{98}$$

where a and b are positive constants whose precise values we are no longer interested in.

We shall now analyze these equations. First observe that, besides the Gaussian fixed point at the origin, there are no other points where both derivatives vanish. Next, the equation for u_0 is readily integrated to give

$$u_0(t)=\frac{u_0(0)}{1+bu_0(0)t}. \tag{99}$$

This means that if we start with a positive coupling $u_0(0)$ and renormalize, the effective coupling renormalizes to zero as $1/t=1/[\ln(\Lambda_0/\Lambda)]$. One says u_0 is *marginally irrelevant*. In the case of bosons a negative u_0 is unphysical, since the functional integral over fields will then diverge for large fields. In some fermion problems one gets the same equation, and negative u_0 is allowed. In that case the coupling is *marginally relevant* and grows. The above equation, derived for weak coupling, will soon have to be abandoned, in contrast to the positive u_0 case, where it gets more and more reliable at larger and larger t. Notice that the fate of marginal couplings (unlike relevant or irrelevant couplings) depends on the sign.

The statement that u_0 is marginally irrelevant at the

Gaussian fixed point needs to be understood properly. In particular, it does *not* mean that if we add a small positive u_0 to the Gaussian fixed point, we shall renormalize back to the Gaussian fixed point. This is because the small u_0 will generate an r_0, and that will quickly grow under renormalization. What is true is that ultimately u_0 will decrease to zero, but r_0 can be large. In fact, to achieve a flow to the Gaussian fixed point, we must start with a particular combination of r_0 and u_0 which describes the critical surface. All this comes out of Eq. (97) for r_0, which is integrated to give

$$r_0(t)=e^{2t}\left[r_0(0)+\int_0^t e^{-2t'}\frac{au_0(0)}{1+bu_0(0)t}dt'\right] . \quad (100)$$

Let us consider large t. Typically r_0 will flow to infinity exponentially fast due to the exponential prefactor, unless we choose r_0 such that the object in brackets vanishes:

$$r_0(0)+\int_0^\infty e^{-2t'}\frac{au_0(0)}{1+bu_0(0)t}dt'=0 , \quad (101)$$

which, for very small $u_0(0)$, translates to

$$r_0(0)=-\frac{au_0(0)}{2} , \quad (102)$$

which defines the critical surface (a line in this case) in the r_0-u_0 plane. Any point on this line approaches the origin as follows:

$$u_0(t)\simeq a/t , \quad (103)$$

$$r_0(t)\simeq e^{2t}\left[-\int_t^\infty e^{-2t'}\frac{au_0(0)}{1+bu_0(0)t}dt'\right]\simeq-\frac{a}{2bt} . \quad (104)$$

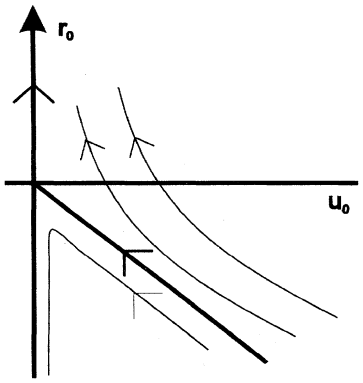

FIG. 3. One-loop flow of the quadratic (r_0) and quartic (u_0) couplings in the scalar field in $d=4$. Notice that, although u_0 is irrelevant, a point on the u_0 axis does not flow to the Gaussian fixed point at the origin. For this to happen, we must tune the parameters to lie on the critical surface, shown by the arrow flowing into the origin.

TABLE II. Fixed-point couplings and flows: a summary.

Item	Comments and relationships
ϕ,ϕ^*	Complex scalar field and its conjugate.
$\langle\phi^*(k_2)\phi(k_1)\rangle\equiv\langle\bar{2}1\rangle$	Two-point function.
$G(k)=1/k^2$	Propagator. $\langle\phi^*(k)\phi(k')\rangle=(2\pi)^4\delta^4(k-k')G(k)$.
Wick's Theorem	Gives N-point functions in terms of 2-point function in Gaussian model.
$\langle\bar{4}\bar{3}21\rangle=\langle\bar{4}2\rangle\langle\bar{3}1\rangle+\langle\bar{4}1\rangle\langle\bar{3}2\rangle$	Four-point function evaluated using Wick's theorem.
$S_0=-\int_\Lambda[d^4k/(2\pi)^4]\phi^*(k)k^2\phi(k)$	Gaussian model action with cutoff Λ.
$\delta S=-\int_0^\Lambda[d^4k/(2\pi)^4]\phi^*(k)r(k)\phi(k)$	Quadratic perturbation.
$r(k)=r_0+r_2k^2+\cdots$	Quadratic coupling function.[a]
$\delta S=-(1/2!2!)\int_\Lambda\phi^*(4)\phi^*(3)\phi(2)\phi(1)u(4321)$	Quartic perturbation in schematic form.
$u(4321)=u_0+O(k)$	Quartic coupling function. $u(4321)=u(3421)=u(4312)$.[b]
RG action	Reduce Λ by s by integrating out fast modes.
	Rewrite result in terms of $\phi'(k')=s^{-3}\phi(k)$, where $k'=sk$.
ζ	Field rescaling factors defined by $\phi'(k')=\zeta\phi(k)$.
	Chosen to make Gaussian action the fixed point.
Cumulant expansion	$\langle e^\Omega\rangle=e^{\langle\Omega\rangle+(\langle\Omega^2\rangle-\langle\Omega\rangle^2)/2+\cdots}$
Tree graphs	Graphs with no closed loops.
Loop graphs	Graphs in which there are closed loops.
	One works to a given number of loops.
Connected graphs	Graphs in which there are no disjoint parts.
Tree-level RG	Calculation with zero loops in Feynman diagrams.[c]
RG flow to one loop	$dr_0/dt=2r_0+au_0$ and $du_0/dt=-bu_0^2$, $a,b>0$.

[a]Only r_0 is relevant; r_2 is marginal, and the rest are irrelevant.
[b]Only u_0 is marginal; higher Taylor coefficients are irrelevant.
[c]Reduced to ignoring fast modes and reexpressing the perturbation in terms of new momenta and fields.

Figure 3 depicts the state of affairs. Table II summarizes the results for the Gaussian fixed point and its leading perturbations.

The analysis of couplings with more powers of the fields is similar. All of them are irrelevant even at the tree level, and higher loops cannot change that. For example, the constant part of the $(\phi^*\phi)^3$ coupling falls like $1/s^2$.

D. The field theory approach to the β function

We just derived the flows in the modern approach, which is intuitively very appealing and consists of integrating out fast modes. We shall now rederive the one-loop flow of u_0 the old way, where the aim is to banish all cutoff dependence from physical quantities. The two approaches will then be compared and contrasted. The reason we even bring up the field theory method is that at higher loops it is more tractable than the modern approach. In the appendices we perform two calculations involving interacting fermions for which the field theory method proves more convenient. The present discussion will be rather succinct, and readers new to diagrams will have to work that much harder.

Consider a field theory with two coupling constants: a mass term r_0, a quartic coupling u_0, and a cutoff Λ. The physical quantity we wish to hold fixed is $\Gamma(k_4\cdots k_1)$, called the *irreducible four-point vertex* or *four-point function*. (Arrows on vectors will be suppressed.) It is defined as follows. Let us define the action of a massive free field S_0,

$$S_0=-\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\phi^*(k)(k^2+r_0)\phi(k)\ . \tag{105}$$

Consider now (suppressing the momentum integration measure for variables labeled 5–8 in the quartic coupling)

$$-\langle\phi^*(k_4)\phi^*(k_3)\phi(k_2)\phi(k_1)\rangle$$

$$=\frac{\int[d\phi^*d\phi][-\phi^*(k_4)\phi^*(k_3)\phi(k_2)\phi(k_1)]e^{S_0}\exp[-(u_0/2!2!)\int_\Lambda\phi^*(k_8)\phi^*(k_7)\phi(k_6)\phi(k_5)]}{\int[d\phi^*d\phi]e^{S_0}\exp[-(u_0/2!2!)\int_\Lambda\phi^*(k_8)\phi^*(k_7)\phi(k_6)\phi(k_5)]}$$

$$\equiv\frac{\langle(-\phi^*(k_4)\phi^*(k_3)\phi(k_2)\phi(k_1))\exp[-(u_0/2!2!)\int_\Lambda\phi^*(k_8)\phi^*(k_7)\phi(k_6)\phi(k_5)]\rangle_0}{\langle\exp[-(u_0/2!2!)\int_\Lambda\phi^*(k_8)\phi^*(k_7)\phi(k_6)\phi(k_5)]\rangle_0} \tag{106}$$

where, in going to the last equation, we have multiplied and divided by the partition function with action S_0, and $\langle\ \rangle_0$ stands for averages with respect to this measure. Notice that all momentum integrals go from 0 to Λ. This is because we are not eliminating modes, we are carrying out a calculation of some correlation function in a given theory. We now calculate the answer in a power series in u_0 by expanding the exponential. We then throw out all disconnected diagrams (diagrams in which some lines are not connected to the others) and delete the four propagators that link the external fields (whose momenta are labeled 1–4) to the vertices that come from the exponential, and the δ function for overall momentum conservation. This defines $\Gamma(4321)$, which is the object we want to be cutoff independent. In field theory, $\Gamma(4321)$ is the scattering amplitude for the process in which $1+2\rightarrow 3+4$, and is a measure of the interaction between particles. In the Gaussian model it will vanish, since all diagrams will be disconnected, the disconnected diagrams describing the independent propagation of noninteracting particles.

Let us now calculate Γ to order u_0.

If we expand the exponential in the numerator to first order, we get a connected piece in which the external fields numbered 1 to 4 get paired with the quartic interaction fields numbered 5 to 8. The factorials get neutralized by the number of ways to pair, the propagators get dropped, and the net result is that to this order

$$\Gamma(4321)=u_0\ . \tag{107}$$

(The denominator is set equal to unity since expanding it to order u_0 will change the answer to order u_0^2.) The reader new to this subject is very strongly urged to carry out the steps using Wick's theorem and paying attention to the combinatorics. Since to lowest order in perturbation theory $\Gamma(4321)=u(4321)$, we shall sometimes refer to u as the amplitude for scattering. To be exact, it is just a coupling constant in the theory, which equals Γ in the weak-coupling limit. It does, however, have all the (anti)symmetries of Γ under the exchange of momentum labels of the external (fermions) bosons.

The above answer for Γ is clearly cutoff independent, and one may choose u_0 to match scattering experiments. Notice that it is also independent of external momenta.

To next order we must expand the numerator to order u_0^2 and the denominator to order u_0 (since Γ starts out at order u_0 in the numerator.) The Feynman diagrams are exactly as before except for the fact that all loop momenta go up to the cutoff. The result is

$$\Gamma(4321)=u_0-u_0^2\left[\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+r_0)(|k+k_3-k_1|^2+r_0)}+\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+r_0)(|k+k_4-k_1|^2+r_0)}\right.$$
$$\left.+\frac{1}{2}\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+r_0)(|-k+k_2+k_1|^2+r_0)}\right]. \tag{108}$$

Let us now demand that $\Gamma(0000)$ be independent of cutoff as the latter goes to infinity. In this limit we find

$$\Gamma(0000)=u_0-u_0^2\frac{5}{32\pi^2}\ln\frac{\Lambda^2}{r_0}. \tag{109}$$

Let us now act on both sides with $d/dt=-\Lambda d/d\Lambda$ and demand that they vanish. This gives us

$$0=\frac{du_0}{dt}+\frac{5}{16\pi^2}u_0^2 \tag{110}$$

from which we find

$$\beta(u_0)=\frac{du_0}{dt}=-\frac{5}{16\pi^2}u_0^2. \tag{111}$$

Two points need clarification here. First: why did we not take the implicit t derivative of the u_0^2 term in Eq. (109)? The answer is that β is of second order in the coupling, and this will give a third-order term. Next one may wonder about Γ with all k's not equal to zero. Will they also be cutoff independent if we choose $u_0(\Lambda)$ as above to make $\Gamma(0000)$ cutoff independent? The answer is yes. If we expand the above integrals in the external momenta, the integrals will become convergent and cutoff independent if we send the cutoff to infinity. Thus the external momenta must be much smaller than the cutoff for the field theory renormalization to work. (If we want the physics to be cutoff independent for external momenta comparable to Λ, we shall need to introduce new couplings besides u_0.)

In the same way one can derive the flow for r_0 by demanding that the pole in the full propagator (the two-point function in the theory with $u_0\neq 0$) have a certain cutoff-independent location.

Suppose we add just the quartic coupling but no mass term r_0 to the Gaussian model. Then we shall find that $\Gamma(0000)$ has an *infrared* logarithmic divergence [sending r_0 to zero in Eq. (109)]. This is a physical divergence in a massless theory, analogous to the infinite cross section for Rutherford scattering in electrodynamics. However the β function, which involves the derivative with respect to the upper limit of momentum integration, is still well defined and has the same value quoted above. In the modern approach, even if $r_0=0$, we shall never see any infrared divergence in the calculation of the β function, since the loop momentum will now go from Λ/s to Λ. This was the meaning of the earlier statement that mode elimination does not produce singularities in the parameters that appear in the effective action because the flow is given by integrals that are well behaved at both ends.

Although the two methods gave the same answer, this is a fact that needs some explanation, since the methods are very different. In the modern approach a change in cutoff is compensated by a change in an infinite number of couplings, while in the latter one tries to compensate by changing just r_0 and u_0. How can this be possible? The answer is that in the field theory approach one always sends the cutoff to infinity (or equivalently looks at correlation functions with external momenta very small compared to the cutoff), while in the modern approach we can ask for quantities involving momenta right up to the cutoff. If in the modern approach we limit ourselves to momenta much smaller than the cutoff, we could trade the complicated Hamiltonian for a simpler one at low momenta, dominated by a few marginal and relevant couplings.

At a graphical level there are differences in the range of integration in the loop graphs that contribute to the flow. In the modern approach we demand that each internal line lie in the shell of width $d\Lambda$ near the cutoff. In the field theory approach, where we take the Λ derivative of momentum integrals going up to Λ, the answer is a sum of terms in which one of the internal propagator momenta is at the cutoff and the others go up to the cutoff. *In our flow equation for u_0 this difference was suppressed.* This was because we argued that only the lowest term in the Taylor series for the coupling was marginal and the rest were irrelevant, allowing us to set all external momenta equal to zero. This meant that if one line in the loop was at the cutoff, the other being either equal to it (ZS and ZS') or opposite to it (BCS) also had to be at the cutoff. *Thus both lines were at the cutoff in both approaches to the flow.* Had we been interested in the renormalization of irrelevant operators, we would have had to consider nonzero external momenta, and the two schemes would have yielded different answers.

Notice that the two schemes do not have to give the same flows, they just have to give the same physics (at momenta much smaller than the cutoff). The bookkeeping can be very different. Consider a more general graph in the field theory approach, with many internal lines and four external lines, so that it contributes to the renormalization of the four-point coupling. Some of these internal momenta may be at the cutoff and the rest below it. Such a graph is forbidden in the modern approach. The effect of these graphs (with slow and fast momenta in the loops) will appear as follows in the modern approach. First all internal lines (propagators) with slow momenta are snipped, and the dangling lines are made into external lines. This graph is then used to renormalize a higher point function with that many more external lines, say six lines in all if two new external lines were produced by

snipping an internal line. Suppose we now stop renormalizing and compute an object, say the four-point vertex of slow modes, using these couplings. The answer will be given as integrals over slow momenta. The six-point coupling that was generated by snipping will contribute to the four-point function when the two lines that got snipped get joined again. In the field theory approach such a contribution would already be sitting inside the effective four-point coupling which got renormalized by graphs with slow and fast lines in the loop.

III. PATH INTEGRALS FOR FERMIONS

In this section the reader is introduced to the path-integral representation of fermion problems. Some elementary problems involving dynamics and thermodynamics will first be solved by operator methods and then the same results will be rederived using the path-integral methods reviewed here. For a more detailed treatment, the reader is referred to the standard references (Berezin, 1966; Schwinger, 1970; Negele and Orland, 1988; Itzykson and Drouffe, 1989).

A. The fermionic oscillator: dynamics and thermodynamics via operators

Let Ψ and $\Psi^\dagger$ be two fermionic operators obeying *anticommutation* relations:

$$\{\Psi^\dagger,\Psi\}=\Psi^\dagger\Psi+\Psi\Psi^\dagger=1 , \tag{112}$$

$$\{\Psi,\Psi\}=\{\Psi^\dagger,\Psi^\dagger\}=0 . \tag{113}$$

Note that the last equation tells us

$$\Psi^{\dagger 2}=\Psi^2=0 . \tag{114}$$

This equation will be used all the time without explicit warning. The number operator

$$N=\Psi^\dagger\Psi \tag{115}$$

obeys

$$N^2=\Psi^\dagger\Psi\Psi^\dagger\Psi=\Psi^\dagger(1-\Psi^\dagger\Psi)\Psi=\Psi^\dagger\Psi=N . \tag{116}$$

Thus the eigenvalues of N can only be 0 or 1. The corresponding normalized eigenstates obey

$$N|0\rangle=0|0\rangle , \tag{117}$$

$$N|1\rangle=1|1\rangle . \tag{118}$$

We shall now prove that

$$\Psi^\dagger|0\rangle=|1\rangle , \tag{119}$$

$$\Psi|1\rangle=|0\rangle . \tag{120}$$

As for the first,

$$N\Psi^\dagger|0\rangle=\Psi^\dagger\Psi\Psi^\dagger|0\rangle=\Psi^\dagger(1-\Psi^\dagger\Psi)|0\rangle=\Psi^\dagger|0\rangle , \tag{121}$$

which shows that $\Psi^\dagger|0\rangle$ has $N=1$. Its norm is unity:

$$\|\Psi^\dagger|0\rangle\|^2=\langle 0|\Psi\Psi^\dagger|0\rangle=\langle 0|(1-\Psi^\dagger\Psi)|0\rangle=\langle 0|0\rangle=1 . \tag{122}$$

It can be similarly shown that $\Psi|1\rangle=|0\rangle$ after first verifying that $\Psi|1\rangle$ is not a null vector, that it has unit norm.

There are no other vectors in the Hilbert space: any attempts to produce more states are thwarted by $\Psi^2=\Psi^{\dagger 2}=0$.

Consider now a Fermi oscillator with Hamiltonian

$$H_0=\Omega_0\Psi^\dagger\Psi , \tag{123}$$

whose eigenvalues are clearly 0 and Ω_0. We shall work not with H_0 but with

$$H=H_0-\mu N , \tag{124}$$

where μ is the *chemical potential.* For the oscillator, since

$$H=(\Omega_0-\mu)\Psi^\dagger\Psi , \tag{125}$$

this merely amounts to measuring all energies relative to the chemical potential.[8]

Let us now turn to thermodynamics. The grand partition function is

$$Z=\mathrm{Tr}\,e^{-\beta(H_0-\mu N)}=e^{-\beta A(\mu,\beta)} , \tag{126}$$

where the trace is over any complete set of eigenstates, β is the inverse temperature $1/T$, and A is the free energy. The latter is clearly a function of μ and β, and its differential is

$$dA=-\langle N\rangle d\mu-S\,dT , \tag{127}$$

where S is the entropy and $\langle N\rangle$ stands for the average particle number. Let us verify that

$$\langle N\rangle=-\frac{\partial A}{\partial\mu} \tag{128}$$

as follows:

$$-\frac{\partial A}{\partial\mu}=\frac{\partial}{\partial\mu}\frac{\ln Z}{\beta} \tag{129}$$

$$=\frac{1}{\beta}\frac{\mathrm{Tr}\beta N e^{-\beta(H_0-\mu N)}}{Z} \tag{130}$$

$$=\frac{1}{\beta}\langle\beta N\rangle=\langle N\rangle . \tag{131}$$

The free energy A is the double Legendre transform of the internal energy $E(S,\langle N\rangle)$:

$$A(\mu,\beta)=E(S,\langle N\rangle)-ST-\mu\langle N\rangle . \tag{132}$$

Thus E must equal $\langle H_0\rangle$. This is indeed so:

$$E=A-T\frac{\partial A}{\partial T}-\mu\frac{\partial A}{\partial\mu} \tag{133}$$

[8]The eigenvalues of H are $T=0$ free energies rather than energies. We shall, however, often refer to H as the Hamiltonian.

$$= A + \beta \frac{\partial A}{\partial \beta} - \mu \frac{\partial A}{\partial \mu} \tag{134}$$

$$= \frac{\partial A \beta}{\partial \beta} - \mu \frac{\partial A}{\partial \mu} \tag{135}$$

$$= \frac{\partial(-\ln Z)}{\partial \beta} + \frac{\mu}{\beta} \frac{\partial \ln Z}{\partial \mu} \tag{136}$$

$$= \langle H_0 \rangle , \tag{137}$$

where the steps leading to the last line are left to the reader. It also follows from the definition of the Legendre transform that

$$\mu = \frac{\partial E}{\partial \langle N \rangle} , \tag{138}$$

so that μ is the (minimum) energy needed to add an extra particle.

The partition function of the Fermi oscillator is easily found (by doing the trace over eigenstates of N) to be

$$Z = 1 + e^{-\beta(\Omega_0 - \mu)} , \tag{139}$$

from which it follows that

$$A = -\frac{1}{\beta} \ln(1 + e^{-\beta(\Omega_0 - \mu)}) , \tag{140}$$

which in turn implies

$$\langle N \rangle = \frac{1}{1 + e^{\beta(\Omega_0 - \mu)}} . \tag{141}$$

We shall be interested in the limit $\beta \to \infty$, in which case

$$\langle N \rangle = \theta(\mu - \Omega_0) , \tag{142}$$

which means the fermion is present if its energy is negative (relative to the chemical potential) and absent if it is positive. This is to be expected, since at $T=0$, $A = E - \mu \langle N \rangle$, and minimizing the free energy is the same as minimizing $\langle H \rangle$.

We now consider the dynamics. From the Schrödinger operators we can form Heisenberg operators:

$$\Psi(t) = e^{iHt} \Psi(0) e^{-iHt} = \Psi(0) e^{-i(\Omega_0 - \mu)t} , \tag{143}$$

$$\Psi^\dagger(t) = e^{iHt} \Psi^\dagger(0) e^{-iHt} = \Psi^\dagger(0) e^{i(\Omega_0 - \mu)t} . \tag{144}$$

We shall study *imaginary-time quantum mechanics*, for which the time evolution operator is

$$U(\tau) = e^{-H\tau} \tag{145}$$

and in which

$$\Psi(\tau) = \Psi(0) e^{-(\Omega_0 - \mu)\tau} , \tag{146}$$

$$\Psi^\dagger(\tau) = \Psi^\dagger(0) e^{(\Omega_0 - \mu)\tau} . \tag{147}$$

Note that, despite the notation, $\Psi(\tau)$ and $\Psi^\dagger(\tau)$ are not adjoints except at $\tau = 0$ owing to the fact that $U(\tau)$ is not unitary.

Readers not familiar with imaginary-time quantum mechanics merely have to observe how the functional formalism reproduces the results of the operator formalism. They may also wish to learn about imaginary-time quantum mechanics using this simple example.

Next consider the *time-ordering symbol* T whose action on a pair of fermionic Heisenberg operators is

$$T(\Psi(\tau)\Psi^\dagger(0)) = \Psi(\tau)\Psi^\dagger(0), \quad \tau > 0 , \tag{148}$$

$$= -\Psi^\dagger(0)\Psi(\tau), \quad \tau < 0 . \tag{149}$$

Note that

$$N = \lim_{\tau \to 0^-} -T(\Psi(\tau)\Psi^\dagger(0)) . \tag{150}$$

In field theory and many-body physics one is interested in the Green's function,

$$G(\tau) = \langle T(\Psi(\tau)\Psi^\dagger(0)) \rangle , \tag{151}$$

where $\langle\ \rangle$ denotes the average with respect to Z. For our problem we find

$$G(\tau) = \frac{\langle 0|T(\Psi(\tau)\Psi^\dagger(0))|0\rangle + \langle 1|T(\Psi(\tau)\Psi^\dagger(0))|1\rangle e^{-\beta(\Omega_0 - \mu)}}{1 + e^{-\beta(\Omega_0 - \mu)}} \tag{152}$$

$$= \frac{\theta(\tau) e^{-(\Omega_0 - \mu)\tau} - \theta(-\tau) e^{-(\Omega_0 - \mu)(\tau + \beta)}}{1 + e^{-\beta(\Omega_0 - \mu)}} . \tag{153}$$

In the zero-temperature limit this reduces to

$$G(\tau) = \theta(\tau) e^{-(\Omega_0 - \mu)\tau}, \quad \mu < \Omega_0 \tag{154}$$

$$= -\theta(-\tau) e^{-(\Omega_0 - \mu)\tau}, \quad \mu > \Omega_0 . \tag{155}$$

Let us define the pair of transforms

$$G(\omega) = \int_{-\infty}^{\infty} G(\tau) e^{i\omega\tau} d\tau , \tag{156}$$

$$G(\tau) = \int_{-\infty}^{\infty} G(\omega) e^{-i\omega\tau} \frac{d\omega}{2\pi} . \tag{157}$$

We find that

$$G(\omega) = \frac{1}{\Omega_0 - \mu - i\omega} , \tag{158}$$

independent of which of Ω_0 or μ is greater.

Let us calculate $\langle N \rangle$ using Eq. (150) and the above Green's function:

$$\langle N \rangle = -G(0^-) \tag{159}$$

$$= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{i\omega 0^+}}{i\omega - \mu - \Omega_0} \tag{160}$$

$$= \theta(\mu - \Omega_0) , \tag{161}$$

where in the last step we have argued that, unless the θ function is satisfied, the contour may be closed in the upper half plane (as dictated by the exponential) which is free of singularities.

Consider finally a toy "Hubbard model" with two fermions and a repulsive interaction U,

$$H_0=\Omega_0(\Psi_1^\dagger\Psi_1+\Psi_2^\dagger\Psi_2)+U\Psi_1^\dagger\Psi_1\Psi_2^\dagger\Psi_2 \tag{162}$$

$$=\Omega_0 N+\frac{U}{2}N(N-1) \tag{163}$$

where

$$N=N_1+N_2=\Psi_1^\dagger\Psi_1+\Psi_2^\dagger\Psi_2\ . \tag{164}$$

(Readers wishing to fill in the missing steps should note that they have to use $N_1^2=N_1$ and likewise for N_2. In any event, they should check the correctness of the final result for various choices of N_1 and N_2.) Each fermion has the usual anticommutator of unity with its adjoint and anticommutes with everything else including all members of the other set.

By summing over the eigenstates of N_1 and N_2, we obtain

$$Z=1+2e^{-\beta(\Omega_0-\mu)}+e^{-\beta(2(\Omega_0-\mu)+U)}\ , \tag{165}$$

where the factor of 2 in the middle term reflects the degeneracy of the one-fermion states. From the above, we obtain by differentiation

$$\langle N\rangle=\lim_{\beta\to\infty}\frac{2(1+e^{\beta(\mu-\Omega_0-U)})}{e^{-\beta(\mu-\Omega_0)}+2+e^{\beta(\mu-\Omega_0-U)}} \tag{166}$$

$$=0\quad \mu<\Omega_0 \tag{167}$$

$$=1\quad \Omega_0<\mu<\Omega_0+U \tag{168}$$

$$=2\quad \mu>\Omega_0+U\ . \tag{169}$$

Table III summarizes results from this subsection. In the next subsection they will be rederived using path integrals.

B. Fermion coherent states

In this section we shall be using *Grassmann numbers.* Here are the rules for manipulating them:

All Grassmann numbers anticommute with each other and with all fermionic operators.

As a result of the above, the square of any Grassmann number is zero and the product of an even number of Grassmann numbers will commute with anything. Likewise any Grassmann number will commute with an even number of fermion operators such as $N=\Psi^\dagger\Psi$. When a Grassmann number is taken through a ket or bra containing an even (odd) number of fermions it will not (will) change sign.

One should not associate a numerical value with Grassmann numbers. There are no large or small Grassmann numbers. All you will need are the above definitions.

Consider the state

$$|\psi\rangle=|0\rangle-\psi|1\rangle\ , \tag{170}$$

where ψ is a Grassmann number. This state, called a *fermion coherent state,* is an eigenstate of Ψ with eigenvalue ψ,

TABLE III. Summary of the fermion oscillators.

Symbol	Comments or definitions
$\Psi,\Psi^\dagger$	Fermion destruction and creation operators. $\{\Psi^\dagger,\Psi\}=\Psi^\dagger\Psi+\Psi\Psi^\dagger=1,\Psi^{\dagger 2}=\Psi^2=0$.
$N=\Psi^\dagger\Psi$	Number operator. $N=0$ or 1.
$\|0\rangle$	State with $N=0$.
$\|1\rangle$	State with $N=1$.
$H_0=\Omega_0\Psi^\dagger\Psi$	Oscillator Hamiltonian.
μ	Chemical potential.
$\beta=1/T$	Inverse temperature.
$H=H_0-\mu N$	Free-energy operator. Also called Hamiltonian.
$Z=\mathrm{Tr}e^{-\beta(H_0-\mu N)}$	Grand partition function Ω.
$\langle\Omega\rangle=\mathrm{Tr}\Omega e^{-\beta(H_0-\mu N)}/\mathrm{Tr}e^{-\beta(H_0-\mu N)}$	Average of operator Ω.
A	Free energy. Defined by $Z=e^{-\beta A(\mu,\beta)}$.
"Hubbard model"	$H_0=\Omega_0 N+(U/2)N(N-1)$
$\langle N\rangle$	Average occupation.[a]
$T(\Psi(\tau)\Psi^\dagger(0))$	$\theta(\tau)\Psi(\tau)\Psi^\dagger(0)-\theta(-\tau)\Psi^\dagger(0)\Psi(\tau)$.
$G(\tau)$	$\langle T(\Psi(\tau)\Psi^\dagger(0))\rangle$. The Green's function.
$-G(0^-)$	N
$G(\omega)$	$\int_{-\infty}^{\infty}G(\tau)e^{i\omega\tau}d\tau=1/(\Omega_0-\mu-i\omega)$
$G(\tau)$	$\int_{-\infty}^{\infty}G(\omega)e^{-i\omega\tau}d\omega/2\pi$

[a]At $\beta\to\infty$ this becomes $\theta(\mu-\Omega_0)$ for single oscillator, $\theta(\mu-\Omega_0)+\theta(\mu-\Omega_0-U)$ for "Hubbard model."

$$\Psi|\psi\rangle=\psi|\psi\rangle \ , \tag{171}$$

as is readily verified:

$$\Psi|\psi\rangle=\Psi|0\rangle-\Psi\psi|1\rangle \tag{172}$$

$$=0+\psi\Psi|1\rangle \tag{173}$$

$$=\psi|0\rangle \tag{174}$$

$$=\psi(|0\rangle-\psi|1\rangle) \tag{175}$$

where we have appealed to the fact that ψ anticommutes with Ψ and that $\psi^2=0$. If we act on both sides of Eq. (171) with Ψ, the left vanishes due to $\Psi^2=0$ and the right due to $\psi^2=0$.

It may be similarly verified that

$$\langle\bar{\psi}|\Psi^\dagger=\langle\bar{\psi}|\bar{\psi} \tag{176}$$

where

$$\langle\bar{\psi}|=\langle 0|-\langle 1|\bar{\psi}=\langle 0|+\bar{\psi}\langle 1| \ . \tag{177}$$

Please note two points. First, the coherent-state vectors are not from the usual complex vector space, since they are linear combinations with Grassmann coefficients. Second, $\bar{\psi}$ is not in any sense the complex conjugate of ψ, and $\langle\bar{\psi}|$ is not the adjoint of $|\psi\rangle$. You should therefore be prepared to see a change of Grassmann variables in which ψ and $\bar{\psi}$ undergo totally unrelated transformations.

The inner product of two coherent states is

$$\langle\bar{\psi}|\psi\rangle=(\langle 0|-\langle 1|\bar{\psi})(|0\rangle-\psi|1\rangle) \tag{178}$$

$$=\langle 0|0\rangle+\langle 1|\bar{\psi}\psi|1\rangle \tag{179}$$

$$=1+\bar{\psi}\psi \tag{180}$$

$$=e^{\bar{\psi}\psi} \ . \tag{181}$$

Any function of a Grassmann variable can be expanded as follows:

$$F(\psi)=F_0+F_1\psi \ , \tag{182}$$

there being no higher powers possible.

We shall now define integrals over Grassmann numbers. These have no geometric significance and are formally defined. We just have to know how to integrate 1 and ψ, since that takes care of all possible functions. Here is the list of integrals:

$$\int\psi\,d\psi=1 \tag{183}$$

$$\int 1\,d\psi=0 \ . \tag{184}$$

That is it! As you can see, selling tables of Grassmann integrals is no way to make a living. There are no limits on these integrals. Integration is assumed to be a linear operation. The differential $d\psi$ is also a Grassmann number. Thus $\int d\psi\,\psi=-1$. The integrals for $\bar{\psi}$ or any other Grassmann variable are identical. A result we shall use often is this:

$$\int\bar{\psi}\psi\,d\psi\,d\bar{\psi}=1 \ . \tag{185}$$

Note that if the differentials or variables come in any other order there can be a change of sign. For example, we shall also invoke the result

$$\int\bar{\psi}\psi\,d\bar{\psi}\,d\psi=-1 \ . \tag{186}$$

Let us now consider some Gaussian integrals. You are urged to show the following:

$$\int e^{-a\bar{\psi}\psi}d\bar{\psi}\,d\psi=a \ , \tag{187}$$

$$\int e^{-\bar{\psi}M\psi}[d\bar{\psi}\,d\psi]=\det M \ , \tag{188}$$

where in the second formula M is a 2×2 matrix, ψ is a column vector with entries ψ_1 and ψ_2, $\bar{\psi}$ is a column vector with entries $\bar{\psi}_1$ and $\bar{\psi}_2$, and $[d\bar{\psi}\,d\psi]=d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2$. This result is true for matrices of any size. To prove these, simply expand the exponential and do the integrals.

Consider next the "averages" over the Gaussian measure:

$$\langle\bar{\psi}\psi\rangle=\frac{\int\bar{\psi}\psi e^{a\bar{\psi}\psi}d\bar{\psi}\,d\psi}{\int e^{a\bar{\psi}\psi}d\bar{\psi}\,d\psi}=\frac{1}{a}=-\langle\psi\bar{\psi}\rangle \ . \tag{189}$$

The proof is straightforward and left as an exercise.

Consider now two sets of Grassmann variables (labeled 1 and 2). The reader should verify that

$$\langle\bar{\psi}_i\psi_j\rangle=\frac{\int\bar{\psi}_i\psi_j e^{a_1\bar{\psi}_1\psi_1+a_2\bar{\psi}_2\psi_2}d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2}{\int e^{a_1\bar{\psi}_1\psi_1+a_2\bar{\psi}_2\psi_2}d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2} \tag{190}$$

$$=\frac{\delta_{ij}}{a_i}\equiv\langle\bar{i}j\rangle \ . \tag{191}$$

We have a Wick's theorem for fermions:

$$\langle\bar{\psi}_i\bar{\psi}_j\psi_k\psi_l\rangle=\frac{\int\bar{\psi}_i\bar{\psi}_j\psi_k\psi_l e^{a_1\bar{\psi}_1\psi_1+a_2\bar{\psi}_2\psi_2}d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2}{\int e^{a_1\bar{\psi}_1\psi_1+a_2\bar{\psi}_2\psi_2}d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2} \tag{192}$$

$$=\frac{\delta_{il}}{a_i}\frac{\delta_{jk}}{a_j}-\frac{\delta_{ik}}{a_i}\frac{\delta_{jl}}{a_j} \tag{193}$$

$$\equiv\langle\bar{i}l\rangle\langle\bar{j}k\rangle-\langle\bar{i}k\rangle\langle\bar{j}l\rangle \ . \tag{194}$$

The reader not familiar with such objects is urged strongly to prove this simple case of Wick's theorem for fermions. Note the strong similarities to the bosonic case. Once again we find that the answer is zero unless each Grassmann is accompanied by its partner. The answer is once again a sum over all possible pairings. The only difference comes from the minus signs, which are determined as follows. We first move each variable until it is next to its partner. In the example above, if $j=k$ and

$i=l$, the middle two Grassmanns are already next to each other and the ends can be brought together without any minus signs, since they are separated by a pair of Grassmanns. This is why the first term is positive in Eq. (194). On the other hand, if $i=k$ and $j=l$, we must move j through k to meet its mate, and this produces a minus sign. When more than four variables are averaged, an obvious generalization holds: pair the fields in all possible ways and put in a minus sign for every time a variable crosses another. Although we did not see it here, the following thing can happen and does happen in a calculation that comes later in this section: the variable and its partner are next to each other, but in the wrong order, with ψ to the left of $\bar{\psi}$. In this case an extra minus sign is needed to rearrange these.

Finally note that Jacobians behave counterintuitively for Grassmann variables. Consider

$$\int a\psi\, d\psi = a \ . \tag{195}$$

In terms of

$$a\psi = \chi \ , \tag{196}$$

$$\int \chi J(\psi/\chi) d\chi = a \ . \tag{197}$$

Assuming J is a constant we pull it out of the integral (with no minus signs, since it involves an even number of Grassmann variables) and use the fact that the integral of χ is unity to obtain

$$J(\psi/\chi) = a \ , \tag{198}$$

while one might have expected the inverse. As an application of this result, the reader may wish to rederive Eq. (187) by making the change of variables from ψ to $\chi = a\psi$. *Note that there is no need to transform $\bar{\psi}$ at the same time.*

We need two more results before we can write down the path integral. The first is the resolution of the identity

$$I = \int |\psi\rangle\langle\bar{\psi}| e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi \ . \tag{199}$$

In the following proof of this result we shall use all the previously described properties and drop terms that are not going to survive integration. (Recall that only $\bar{\psi}\psi = -\psi\bar{\psi}$ has a nonzero integral.)

$$\int |\psi\rangle\langle\bar{\psi}| e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi = \int |\psi\rangle\langle\bar{\psi}| (1-\bar{\psi}\psi) d\bar{\psi}\, d\psi \tag{200}$$

$$= \int (|0\rangle - \psi|1\rangle)(\langle 0| - \langle 1|\bar{\psi})(1-\bar{\psi}\psi) d\bar{\psi}\, d\psi \tag{201}$$

$$= \int (|0\rangle\langle 0| + \psi|1\rangle\langle 1|\bar{\psi})(1-\bar{\psi}\psi) d\bar{\psi}\, d\psi \tag{202}$$

$$= |0\rangle\langle 0| \int (-\bar{\psi}\psi) d\bar{\psi}\, d\psi + |1\rangle\langle 1| \int \psi\bar{\psi}\, d\bar{\psi}\, d\psi \tag{203}$$

$$= I \ . \tag{204}$$

The final result we need is that for any bosonic operator (an operator made of an even number of Fermi operators)

$$\mathrm{Tr}\Omega = \int \langle -\bar{\psi}|\Omega|\psi\rangle e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi \ . \tag{205}$$

The proof is very much like the one just given and is left to the reader.

C. The fermionic path integral

Consider the partition function for a single oscillator:

$$Z = \mathrm{Tr} e^{-\beta(\Omega_0-\mu)\Psi^\dagger\Psi} \tag{206}$$

$$= \int \langle -\bar{\psi}| e^{-\beta(\Omega_0-\mu)\Psi^\dagger\Psi} |\psi\rangle e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi \ . \tag{207}$$

You cannot simply replace $\Psi^\dagger$ and Ψ by $-\bar{\psi}$ and ψ, respectively, in the exponential. This is because when we expand out the exponential not all the Ψ's will be acting to the right on their eigenstates and neither will all $\Psi^\dagger$'s be acting to the left on their eigenstates. (Remember that we are dealing with operators, not Grassmann numbers. The exponential will have an infinite number of terms in its expansion.) We need to convert the exponential to its *normal-ordered form,* in which all the creation operators stand to the left and all the destruction operators to the right. Luckily we can write down the answer by inspection:

$$e^{-\beta(\Omega_0-\mu)\Psi^\dagger\Psi} = 1 + (e^{-\beta(\Omega_0-\mu)} - 1)\Psi^\dagger\Psi \ , \tag{208}$$

whose correctness we can verify by considering the two possible values of $\Psi^\dagger\Psi$. (Alternatively we could expand the exponential and use the fact that $N^k = N$ for any nonzero k.) Now we may write

$$Z = \int \langle -\bar{\psi}| 1 + (e^{-\beta(\Omega_0-\mu)} - 1)\Psi^\dagger\Psi |\psi\rangle e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi \tag{209}$$

$$= \int \langle -\bar{\psi}|\psi\rangle [1 + (e^{-\beta(\Omega_0-\mu)} - 1)(-\bar{\psi}\psi)] e^{-\bar{\psi}\psi} d\bar{\psi}\, d\psi \tag{210}$$

$$= \int [1 - (e^{-\beta(\Omega_0-\mu)} - 1)\bar{\psi}\psi] e^{-2\bar{\psi}\psi} d\bar{\psi}\, d\psi \tag{211}$$

$$= 1 + e^{-\beta(\Omega_0-\mu)} \tag{212}$$

as expected. While this is the right answer, this is not the path-integral approach. As for the latter, the procedure

is this. Consider

$$Z = \mathrm{Tr} e^{-\beta H} , \tag{213}$$

where H is a normal-ordered operator $H(\Psi^\dagger, \Psi)$. We write

$$e^{-\beta H} = \lim_{N\to\infty} (e^{-(\beta/N)H})^N \tag{214}$$

$$= \underbrace{(1-\varepsilon H)\cdots(1-\varepsilon H)}_{N_{\text{times}}}, \quad \varepsilon = \beta/N , \tag{215}$$

and introduce the resolution of the identity $N-1$ times:

$$Z = \int \langle -\bar\psi_1|(1-\varepsilon H)|\psi_{N-1}\rangle e^{-\bar\psi_{N-1}\psi_{N-1}} \langle \bar\psi_{N-1}|(1-\varepsilon H)|\psi_{N-2}\rangle e^{-\bar\psi_{N-2}\psi_{N-2}} \langle \psi_{N-2}| \times \cdots |\psi_2\rangle e^{-\bar\psi_2\psi_2}\langle\bar\psi_2|(1-\varepsilon H)|\psi_1\rangle e^{-\bar\psi_1\psi_1} \prod_{i=1}^{N-1} d\bar\psi_i d\psi_i . \tag{216}$$

Now we may legitimately make the replacement

$$\langle \bar\psi_{i+1}|1-\varepsilon H(\Psi^\dagger,\Psi)|\psi_i\rangle = \langle \bar\psi_{i+1}|1-\varepsilon H(\bar\psi_{i+1},\psi_i)|\psi_i\rangle = e^{\bar\psi_{i+1}\psi_i} e^{-\varepsilon H(\bar\psi_{i+1},\psi_i)} , \tag{217}$$

where in the last step we are anticipating the limit of infinitesimal ε. Let us now *define* an additional pair of variables (not to be integrated over),

$$\bar\psi_N = -\bar\psi_1 , \tag{218}$$

$$\psi_N = -\psi_1 . \tag{219}$$

The first of these equations allows us to replace the leftmost bra in Eqs. (216), $\langle -\bar\psi_1|$, by $\langle \bar\psi_N|$. The reason for introducing ψ_N will follow soon.

Putting together all the factors (including the overlap of coherent states), we end up with

$$Z = \int \prod_{i=1}^{N-1} e^{\bar\psi_{i+1}\psi_i} e^{-\varepsilon H(\bar\psi_{i+1},\psi_i)} e^{-\bar\psi_i\psi_i} d\bar\psi_i d\psi_i \tag{220}$$

$$= \int \prod_{i=1}^{N-1} \exp\left[\frac{(\bar\psi_{i+1}-\bar\psi_i)}{\varepsilon}\psi_i - H(\bar\psi_{i+1},\psi_i)\right]\varepsilon d\bar\psi_i d\psi_i \tag{221}$$

$$\simeq \int e^{\int_0^\beta \bar\psi(\tau)(-\partial/\partial\tau-\Omega_0+\mu)\psi(\tau)d\tau}[d\bar\psi\, d\psi] , \tag{222}$$

where the last step needs some explanation. With all the factors of ε in place we do seem to get the continuum expression in the last formula. However, the notion of replacing differences by derivatives is purely symbolic for Grassmann variables. There is no sense in which $\bar\psi_{i+1}-\bar\psi_i$ is small; in fact, the objects have no numerical values. What this really means here is the following. In a while we shall trade $\psi(\tau)$ for $\psi(\omega)$ related by Fourier transformation. At that stage we shall replace $-\partial/\partial\tau$ by $i\omega$, while the exact answer is $e^{i\omega}-1$. If we do not make this replacement, the Grassmann integral, when evaluated in terms of ordinary numbers, will give exact results for anything one wants to calculate, say the free energy. With this approximation, only quantities insensitive to high frequencies will be given correctly. The free energy will come out wrong, but the correlation functions will be correctly reproduced. (This is because the latter are given by derivatives of the free energy, and these derivatives make the integrals sufficiently insensitive to high frequencies.) Notice also that we are replacing

$$H(\bar\psi_{i+1},\psi_i) = H(\bar\psi(\tau+\varepsilon),\psi(\tau))$$

by $H(\bar\psi(\tau),\psi(\tau))$ in the same spirit.

Now turn to the Fourier expansions alluded to above. Let us write

$$\bar\psi(\tau) = \sum_n \frac{e^{i\omega_n\tau}}{\beta}\bar\psi(\omega) , \tag{223}$$

$$\psi(\tau) = \sum_n \frac{e^{-i\omega_n\tau}}{\beta}\psi(\omega) , \tag{224}$$

where the allowed frequencies are chosen to satisfy the antisymmetric boundary conditions in Eqs. (218) and (219). Thus

$$\omega_n = \frac{(2n+1)\pi}{\beta} , \tag{225}$$

where n is an integer. Note that we have chosen the Fourier expansions as if ψ and $\bar\psi$ were complex conjugates, which they are not. This choice, however, makes the calculations easy.

The inverse transformations are

$$\psi(\omega) = \int_0^\beta \psi(\tau) e^{i\omega_n\tau} d\tau , \tag{226}$$

$$\bar\psi(\omega) = \int_0^\beta \bar\psi(\tau) e^{-i\omega_n\tau} d\tau , \tag{227}$$

where we use the orthogonality property

$$\int_0^\beta e^{i\omega_n\tau} e^{-i\omega_m\tau} d\tau = \frac{e^{i(\omega_n-\omega_m)\beta}-1}{i(\omega_n-\omega_m)} = \beta\delta_{mn} . \tag{228}$$

Performing the Fourier transforms in the action and changing the functional integration variables to $\psi(\omega)$ and $\bar\psi(\omega)$ (the Jacobian is unity) and going to the limit $\beta\to\infty$, which converts sums over discrete frequencies to integrals over a continuous ω, we end up with

$$Z = \int \exp\left[\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\bar\psi(\omega)(i\omega-\Omega_0+\mu)\psi(\omega)\right] \times [d\bar\psi(\omega)d\psi(\omega)] . \tag{229}$$

Although β has disappeared from the picture, it *will appear as* $2\pi\delta(0)$, which we know stands for the total time. (Recall Fermi's "golden rule" calculations.) An example will follow shortly.

Let us first note that, just as in the case of the scalar Gaussian model, the correlation function is related to the integral over just a single pair of variables [Eq. (187)] and is given by

$$\langle \bar{\psi}(\omega_1)\psi(\omega_2)\rangle = \frac{\int \bar{\psi}(\omega_1)\psi(\omega_2)\exp\left[\int_{-\infty}^{\infty}(d\omega/2\pi)\bar{\psi}(\omega)(i\omega-\Omega_0+\mu)\psi(\omega)\right][d\bar{\psi}(\omega)d\psi(\omega)]}{\int d\bar{\psi}(\omega)d\psi(\omega)\exp\left[\int_{-\infty}^{\infty}(d\omega/2\pi)\bar{\psi}(\omega)(i\omega-\Omega_0+\mu)\psi(\omega)\right]} \tag{230}$$

$$= \frac{2\pi\delta(\omega_1-\omega_2)}{i\omega_1-\Omega_0+\mu} . \tag{231}$$

In particular,

$$\langle \bar{\psi}(\omega)\psi(\omega)\rangle = \frac{2\pi\delta(0)}{i\omega-\Omega_0+\mu} = \frac{\beta}{i\omega-\Omega_0+\mu} . \tag{232}$$

Let us now calculate the mean occupation number $\langle N\rangle$:

$$\langle N\rangle = \frac{1}{\beta Z}\frac{\partial Z}{\partial \mu} \tag{233}$$

$$= \frac{1}{\beta}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\langle \bar{\psi}(\omega)\psi(\omega)\rangle \tag{234}$$

$$= \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{i\omega 0^+}}{i\omega-\Omega_0+\mu} \tag{235}$$

$$= \theta(\mu-\Omega_0) , \tag{236}$$

as in the operator approach. Notice that we had to introduce the factor $e^{i\omega 0^+}$ into the ω integral. We understand this as follows. If we had done the calculation using time τ instead of frequency ω, we would have calculated the average of $\Psi^\dagger\Psi$. This would automatically have turned into $\bar{\psi}(\tau+\varepsilon)\psi(\tau)$ when introduced into the path integral, since the coherent state bra to the left of the operator would have come from the next time slice compared to the ket at the right. [Remember how $H(\Psi^\dagger,\Psi)$ turned into $H(\bar{\psi}(i+1)\psi(i))$.] Notice that the integral over ω was not convergent, varying as $d\omega/\omega$. It was therefore sensitive to the high frequencies, and we had to intervene with the factor $e^{i\omega 0^+}$. Later we shall deal with integrals that have two or more powers of ω in the denominator and are hence convergent. We shall not introduce this factor in those cases.

Our final calculation will be the determination of $\langle N\rangle$ for the toy Hubbard model using path integrals. The partition function is

$$Z = \int [d\bar{\psi}_1 d\psi_1 d\bar{\psi}_2 d\psi_2] e^{S_0} e^{U\int \bar{\psi}_1\bar{\psi}_2\psi_1\psi_2} \tag{237}$$

where

$$S_0 = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\sum_{i=1}^{2}\bar{\psi}_i(\omega)(i\omega-\Omega_0+\mu)\psi_i(\omega) , \tag{238}$$

$$U\int \bar{\psi}_1\bar{\psi}_2\psi_1\psi_2 = \int_{-\infty}^{\infty}\prod_{i=1}^{4}\frac{d\omega_i}{2\pi}\bar{\psi}_1(\omega_4)\bar{\psi}_2(\omega_3)\psi_1(\omega_2)\psi_2(\omega_1)2\pi\delta(\omega_4+\omega_3-\omega_2-\omega_1) . \tag{239}$$

Notice that the Hamiltonian already was in normal-ordered form: each creation operator was to the left of *its* destruction operator. This allowed us to replace the operators by their eigenvalues when the coherent states were introduced. For notational uniformity we have further arranged to have all creation operators to the left of all destruction operators. This merely introduces an extra minus sign here, since operators corresponding to different oscillators anticommute.

Prior to calculating $\langle N\rangle$ let us calculate $\langle N_1\rangle$. This is given by

$$N_1 = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle e^{i\omega 0^+} , \tag{240}$$

where $\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle$ stands for the correlation function with the full action and not just S_0. We may, however, express it in terms of averages over the Gaussian measure S_0, using the same trick we used for bosons: we multiply and divide the exact Z by Z_0, the partition function with $U=0$, to obtain

$$\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle = \frac{\langle \bar{\psi}_1(\omega)\psi_1(\omega)e^{U\int \bar{\psi}_1\bar{\psi}_2\psi_1\psi_2}\rangle_0}{\langle e^{U\int \bar{\psi}_1\bar{\psi}_2\psi_1\psi_2}\rangle_0} . \tag{241}$$

In principle the reader has all the information needed to evaluate this expression to any order in perturbation theory. The exponential is to be expanded (in the numerator and denominator) to the desired order and all averages done using Wick's theorem. We shall carry out the calculation to order U to show the details. The result to all orders will then simply be stated, and the skeptics are encouraged to check it to higher orders. The integration measure and delta functions will occasionally be suppressed.

To first order we find

$$\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle = \frac{\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle_0 + U\int \langle \bar{\psi}_1(\omega)\psi_1(\omega)\bar{\psi}_1(\omega_4)\bar{\psi}_2(\omega_3)\psi_1(\omega_2)\psi_2(\omega_1)\rangle_0}{1+U\int \langle \bar{\psi}_1(\omega_4)\bar{\psi}_2(\omega_3)\psi_1(\omega_2)\psi_2(\omega_1)\rangle_0} . \tag{242}$$

In both the numerator and the denominator we have just one pair of fields with label 2; these must obviously be paired, though it takes a sign change to bring them next to each other. The four fields with label 1 can be paired in two ways. One of the ways, in which the external fields $\bar{\psi}_1(\omega)\psi_1(\omega)$ get paired is precisely canceled by the denominator expanded out to order U. This corresponds to the cancellation of disconnected diagrams. Readers new to this concept are urged to verify this. What remains may be written

$$\langle \bar{\psi}_1(\omega)\psi_1(\omega)\rangle = \frac{2\pi\delta(0)}{i\omega-\Omega_0+\mu} + \frac{2\pi\delta(0)}{(i\omega-\Omega_0+\mu)^2} U\int_{-\infty}^{\infty} \frac{d\Omega_1}{2\pi} \frac{e^{i\omega_1 0^+}}{i\omega_1-\Omega_0+\mu} \tag{243}$$

$$= \frac{2\pi\delta(0)}{i\omega-\Omega_0-\mu-U\langle N_2\rangle} . \tag{244}$$

In going to the last step we have taken two terms of the power series in U, assumed that they represent a geometric series, and summed the series. This result is undoubtedly correct to the order we are working in. We have also replaced the integral over ω_1 with $\langle N_2\rangle$ which is also good to this order. It turns out that both these approximations are in fact exactly what we would get if we went to all orders, as will be explained shortly. Let us accept this for the present and see what follows. Using the above correlation function in the formula for $\langle N_1\rangle$, Eq. (240), we obtain

$$\langle N_1\rangle = \theta[\mu-\Omega_0-U\langle N_2\rangle] . \tag{245}$$

Obviously we can similarly derive another equation,

$$\langle N_2\rangle = \theta[\mu-\Omega_0-U\langle N_1\rangle] . \tag{246}$$

Let us explore these equations for various cases. First, if $\mu<\Omega_0$, neither θ function can be satisfied, since $\langle N_i\rangle \geq 0$. Thus we get $\langle N\rangle = 0$ for this case as before. Likewise if $\mu>\Omega_0+U$, both θ functions will be satisfied, since $\langle N_i\rangle \leq 1$. This gives us $\langle N\rangle = 2$ as before. Finally, consider $\Omega_0<\mu<\Omega_0+U$. Since each $\langle N_i\rangle$ equals a θ function, it can equal only 0 or 1. It is readily seen from these equations that the only two consistent choices are $\langle N_1\rangle = 1$, $\langle N_2\rangle = 0$ and vice versa, once again in agreement with the operator solution.

Now for the higher terms in the expansion. These are best seen in diagrammatic terms. Consider Fig. 4(a). The heavy line stands for the full propagator and the thin line for the one computed in the Gaussian measure. To order U we kept the two diagrams shown, and these correspond to the expressions in Eq. (243). The disconnected diagram that got canceled by the denominator is shown in Fig. 4(b). If we go to higher orders, we can run into either iterates of the one-loop diagram or embellishments of it. The embellishments convert the loop integral over the free propagator of species 2 to the integral over the full propagator, which then reduces to $\langle N_2\rangle$. The iterations produce the remaining terms in the geometric series that was presumed in going from Eq. (243) to the next one. This leaves us with diagrams such as the "sunrise" diagram in Fig. 4(d). (Once again the nomenclature is from field theory.) These diagrams and all the rest vanish in this problem because the corresponding frequency integrals are convergent and have all the poles on the same half plane, allowing us to close the contour the other way.

The generalization of Grassmann integrals to many-body problems is straightforward. The labels 1 and 2 from the toy Hubbard model can run over, say, the modes in the Brillouin zone. The action is once again obtained by replacing the normal-ordered Hamiltonian by the corresponding function of Grassmannian coherent-state labels. As for the coupling functions, just as the coupling in the bosonic $u(4321)$ was symmetric under the exchange of the first or last two labels among themselves, the fermionic couplings will be antisymmetric under such an exchange due to the anticommuting nature of the

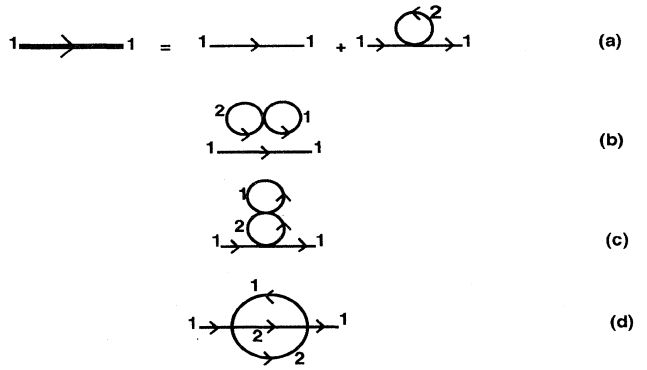

FIG. 4. Graphical representation of the perturbative calculation of the type-1 fermionic propagator in the toy Hubbard model; (a) the two terms we kept in the analysis to one loop; (b) corrections that are grahically disconnected and ignorable, since they get canceled by the partition function that comes in the denominator of all averges; (c) a correction that embellishes the type-2 propagator in the loop we did consider. The effect of such graphs is to turn the free propagator in the loop to the exact propagator for 2, which in turn means that the integral over the loop equals the exact density N_2. (d) a two-loop contribution that vanishes upon ω integration. Not shown are iterates of the connected diagrams which are part of the geometric series, which we assumed and summed in the text.

Grassmann variables.

Finally a matter of notation. We shall be switching from upper-case letters Ψ and $\Psi^\dagger$ for fermion operators to lower-case. It will be clear from the context whether we are referring to the operators or the Grassmann variables.

Table IV summarizes the results from the discussion on fermionic path integrals.

IV. MOTIVATION AND WARMUP: RG IN $d=1$

We are now ready to turn to the main topic: application of the RG to interacting nonrelativistic fermions. The best way to explain the method is to deal with specific problems to which it applies. We begin with the problem of charge-density-wave (CDW) formation in a system of spinless fermions at half filling. It has relatively simple kinematics and illustrates the RG approach very nicely. In fact, the methods explained here were originally developed (Shankar, 1991) to deal with this problem in two dimensions. Let us begin with a discussion of the various terms used above in describing the model.

We begin with the justification for the study of spinless fermions, which are admittedly a theorist's construction. As we progress with this paper, it will become apparent that the RG is primarily concerned with the symmetry properties of the Fermi surface of the noninteracting fermions. For example, the superconducting instability for arbitrarily small attraction is due to the invariance of the Fermi surface under time reversal: if $\mathbf{K}$ lies on the surface, so does $-\mathbf{K}$. Likewise the charge-density-wave (CDW) instability on the square lattice, towards a ground

TABLE IV. Fermion path integrals and useful relations. ψ and $\bar{\psi}$ are Grassmann numbers. They anticommute with each other and with fermion creation and destruction operators. Their differentials are also Grassmann numbers.

$$|\psi\rangle = |0\rangle - \psi|1\rangle$$
$$\langle\bar{\psi}| = \langle 0| - \langle 1|\bar{\psi}$$
$$\Psi|\psi\rangle = \psi|\psi\rangle$$
$$\langle\bar{\psi}|\Psi^\dagger = \langle\bar{\psi}|\bar{\psi}$$
$$\langle\bar{\psi}|\psi\rangle = e^{\bar{\psi}\psi}$$
$$\langle\bar{\psi}|H(\Psi^\dagger,\Psi)|\psi\rangle = H(\bar{\psi},\psi)$$
$$\int \psi d\psi = 1$$
$$\int 1 d\psi = 0$$
$$\int \bar{\psi} d\bar{\psi} = 1$$
$$\int 1 d\bar{\psi} = 0$$
$$\int \bar{\psi}\psi d\psi d\bar{\psi} = 1 = -\int \bar{\psi}\psi d\bar{\psi} d\psi$$
$$\int e^{-\bar{\psi}_i M_{ij}\psi_j}\left[\prod_i d\bar{\psi}_i d\psi_i\right] = \det M$$
$$\langle\bar{\psi}_i\psi_j\rangle_0 = \frac{\int \bar{\psi}_i\psi_j e^{\sum_k a_k\bar{\psi}_k\psi_k}\prod_k d\bar{\psi}_k d\psi_k}{\int e^{\sum_k a_k\bar{\psi}_k\psi_k}\prod_k d\bar{\psi}_k d\psi_k} = \frac{\delta_{ij}}{a_i} \equiv \langle\bar{i}j\rangle$$
$$\langle\bar{\psi}_i\bar{\psi}_j\psi_k\psi_l\rangle_0 = \frac{\delta_{il}\delta_{jk}}{a_i a_j} - \frac{\delta_{ik}\delta_{jl}}{a_i a_j} \equiv \langle\bar{i}l\rangle\langle\bar{j}k\rangle - \langle\bar{i}k\rangle\langle\bar{j}l\rangle \text{ (Wick's theorem)}$$
$$I = \int |\psi\rangle\langle\bar{\psi}|e^{-\bar{\psi}\psi} d\bar{\psi} d\psi$$
$$\mathrm{Tr}\Omega = \int d\bar{\psi} d\psi \langle -\bar{\psi}|\Omega|\psi\rangle e^{-\bar{\psi}\psi}$$
$$Z_{\text{oscillator}} = \int [d\bar{\psi}(\tau) d\psi(\tau)] e^{\int_0^\beta \bar{\psi}(\tau)(-\partial/\partial\tau - \Omega_0 + \mu)\psi(\tau) d\tau}$$
$$Z_{\text{oscillator},\beta=\infty} = \int [d\bar{\psi}(\omega) d\psi(\omega)] e^{\int_{-\infty}^{\infty} (d\omega/2\pi)\bar{\psi}(\omega)(i\omega - \Omega_0 + \mu)\psi(\omega)}$$
$$\langle\bar{\psi}(\omega_1)\psi(\omega_2)\rangle = \frac{2\pi\delta(\omega_1 - \omega_2)}{i\omega_1 - \Omega_0 + \mu}$$
$$\langle\bar{\psi}(\omega)\psi(\omega)\rangle = \frac{2\pi\delta(0)}{i\omega - \Omega_0 + \mu}$$
$$2\pi\delta(0) = \beta$$
$$\langle N\rangle = \frac{1}{\beta Z}\frac{\partial Z}{\partial\mu} = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{i\omega 0^+}}{i\omega - \Omega_0 + \mu}$$

state in which the charge density is nonuniform and oscillates between the sublattices, is due to nesting: if $\mathbf{K}$ lies on the Fermi surface, so does $\mathbf{K}+\mathbf{Q}_N$, where $\mathbf{Q}_N$ is a fixed nesting vector. Now these very same properties of the Fermi surface will also destabilize a system of real electrons. The actual nature of the instability can of course be different in the two cases. For example, the nested Fermi surface in $d=2$ will cause electrons to go to an antiferromagnetic state in which the magnetization (rather than charge density) oscillates in magnitude between the two sublattices. Likewise, the time-reversal-invariant Fermi surface will lead to superconductivity, but (Cooper) pairing can take place for any angular momentum, while for spinless fermions only odd orbital angular momentum states are allowed due to antisymmetry requirements. To summarize, spin really is an unessential complication if we are simply trying to understand how the RG works. When comparing theory to experiment, we will of course have to include spin, but this really will be straightforward.

Consider now the requirement of half-filling, which means that the system has half the maximum number allowed by the exclusion principle. For spinless fermions this means one particle for every other site, while for electrons it means one particle per site. Despite this, the two problems will have the same Fermi surface in the noninteracting case. In both cases, the Fermi surface will enclose half the Brillouin zone and have the same shape, decided by the lattice parameters. However, each filled momentum state will carry two electrons (of opposite spin) but just one spinless fermion. In other words, the condition of half filling implies a different particle density in the two cases, but the same Fermi surface.[9]

We start here with the one-dimensional version of the spinless fermion problem for two reasons. First, the one-dimensional problem is very interesting in itself and shows the power of the RG approach. Secondly, as mentioned earlier, due to the fact that the Fermi surface in $d=1$ consists of just two points, the problem resembles quantum field theory with a few coupling constants (rather than a theory with coupling functions) and affords a painless introduction to the use of the RG for fermions. We shall then be better prepared for the $d=2$ version in Sec. VI.

[9]An aside for readers with a different background, say, particle physics, who are troubled by the following: why bother with effects at such a special filling? Is there any chance that a generic system will have filling factor of exactly half, as compared to say 0.51? Yes! Consider a square lattice with an atom at each site. Since each atom contributes an *integer* number of electrons to the conduction band, the filling factor (the ratio of the number of electrons to the maximum allowed number of two per site) is bound to be a half integer or integer. For a more complicated unit cell, or intercalated compounds, one can have other simple fractions like $\frac{1}{4}$.

A. The $d=1$ model: definition and mean-field analysis

Let us consider the following specific Hamiltonian for a spinless fermion system on a $d=1$ lattice labeled by an integer n:

$$H=H_0+H_I \tag{247}$$

$$=-\tfrac{1}{2}\sum_j \psi^+(j+1)\psi(j)+\text{H.c.} + U_0\sum_j [\psi^\dagger(j)\psi(j)-\tfrac{1}{2}][\psi^\dagger(j+1)\psi(j+1)-\tfrac{1}{2}] , \tag{248}$$

where the fields obey

$$\{\psi^\dagger(j),\psi(m)\}=\delta_{mj} , \tag{249}$$

with all other anticommutators vanishing.

The first term represents hopping. The hopping amplitude has been normalized to $\frac{1}{2}$. The second term represents nearest-neighbor repulsion of strength U_0. The role of the $\frac{1}{2}$'s subtracted from the charge densities n_j $(=\psi_j^\dagger\psi_j)$ and n_{j+1} is this. When we open up the brackets, it is readily seen that they represent a chemical potential

$$\mu=U_0 . \tag{250}$$

This happens to be exactly the value needed to maintain half filling in the presence of the repulsion U_0. To see this, make the change $\psi\leftrightarrow\psi^\dagger$ at all sites. This exchanges the site occupation number $n=\psi^\dagger\psi$ with $1-n$ or changes the sign of $n-\frac{1}{2}$. Thus both brackets in the interaction term change sign under this, and their product is unaffected. As for the hopping term, it changes sign under $\psi\leftrightarrow\psi^\dagger$. This can be compensated by changing the sign of $\psi,\psi^\dagger$ on just one sublattice (which preserves the anticommutation rules and does not affect the other term). Thus H is invariant under exchanging particles with holes. This means the ground state (if it is unique) will satisfy $\langle n\rangle=\langle 1-n\rangle$, which in turn means $\langle n\rangle=\frac{1}{2}$. (If there is a degeneracy of the ground state, the result still holds, but takes a little more work to establish.)

Let us understand this model in the extreme limits $U_0=0$ and $U_0=\infty$. As for the first case let us introduce momentum states via

$$\psi(j)=\int_{-\pi}^{\pi}\frac{dK}{2\pi}\psi(K)e^{iKj} \tag{251}$$

and the inverse transform

$$\psi(K)=\sum_j e^{-iKj}\psi(j) . \tag{252}$$

Using $\Sigma_j=2\pi\delta(0)$, we can verify that

$$\{\psi(K),\psi^\dagger(K')\}=2\pi\delta(K-K') . \tag{253}$$

In terms of these operators

$$H_0=\int_{-\pi}^{\pi}\frac{dK}{2\pi}\psi^\dagger(K)\psi(K)E(K)\ , \tag{254}$$

$$E(K)=-\cos K\ . \tag{255}$$

The Fermi sea is obtained by filling all negative energy states, i.e., those with $|K|\le K_F=\pi/2$, which corresponds to half filling. The Fermi surface consists of just two points $|K|=\pm\pi/2$. It is clear that the ground state is a perfect conductor, since we can move a particle just below the Fermi surface to just above it at arbitrarily small energy cost.

Consider now the situation in the other extreme, $U_0=\infty$. We now ignore the hopping term and focus on just the interaction. It is evident that the lowest-energy states are those in which no particle has a neighbor: thus either the A sublattice, consisting of even sites, is occupied or the B sublattice, made up of the odd sites, is occupied. This makes the product $(n_j-\frac{1}{2})(n_{j+1}-\frac{1}{2})$ negative on every bond. These two states, which break the translational symmetry of the lattice are the CDW states. The order parameter, which measures the difference between the mean occupation of the odd and even sites, is maximal (unity). In the CDW state, the system is an insulator. Any excitation of the ground state requires us to move the charge, and this will cost an energy of order U_0. (This is clearest as $U_0\to\infty$.) One expects that, even for large but finite U_0, the symmetry will still be broken, but with a smaller order parameter.

Here is the question we want to answer: *Will the system develop the CDW order and gap for arbitrarily small repulsion, or will it remain a conductor up to some finite* U_0?

We shall use the RG to answer it. But first let us see what a very standard tool, namely, *mean-field theory,* can tell us. In this approach one assumes a CDW order parameter in the ground state and asks if the assumption is self-consistent. The self-consistency check is approximate, as will be explained. Mean-field theory predicts that charge-density waves will set in for any repulsion, however small. Here is a short description of the calculation.

Let us begin with Eq. (248) and make the *ansatz*

$$\langle n_j\rangle=\tfrac{1}{2}+\tfrac{1}{2}(-1)^j\Delta\ , \tag{256}$$

where Δ is the CDW order parameter. We shall now see if the ground-state energy of the system is lowered by a nonzero value of Δ. To this end, we find the ground-state energy as a function of Δ and minimize it and see if the minimum occurs at a nonzero Δ. However, this last step will be done approximately since this is an interacting many-body system. The approximation is the following.

We start with the interaction

$$H=-\tfrac{1}{2}\sum_j\psi^\dagger(j+1)\psi(j)+\text{H.c.}+U_0\sum_j(n_j-\tfrac{1}{2})(n_{j+1}-\tfrac{1}{2}) \tag{257}$$

and make the substitution

$$n_j=\tfrac{1}{2}+\tfrac{1}{2}(-1)^j\Delta+{:}n_j{:}\ , \tag{258}$$

$$n_{j+1}=\tfrac{1}{2}+\tfrac{1}{2}(-1)^{j+1}\Delta+{:}n_{j+1}{:}\ , \tag{259}$$

where the normal-ordered operator $:n_j:$ has no expectation value in the true ground state and represents the fluctuations in number density. Upon making these substitutions and some rearrangements, we find

$$H=-\frac{1}{2}\sum_j\psi^\dagger(j+1)\psi(j)+\text{H.c.} \tag{260}$$

$$+U_0\left[\tfrac{1}{4}\sum_j\Delta^2-\Delta\sum_j(-1)^jn_j\right]+U_0\sum_j{:}n_j{::}n_{j+1}{:}\ . \tag{261}$$

In the mean-field approximation we ignore the last term. The rest of the Hamiltonian is quadratic and solved by Fourier transformation. Due to the factor $(-1)^j$, which multiplies n_j, states with momentum K and $K'=K+\pi$ will mix. The Hamiltonian becomes

$$H=\int_0^\pi\frac{dK}{2\pi}(\psi^\dagger(K),\psi^\dagger(K'))\begin{bmatrix}E(K) & -U_0\Delta\\ -U_0\Delta & E(K')\end{bmatrix}\begin{bmatrix}\psi(K)\\ \psi(K')\end{bmatrix}+U_0 2\pi\delta(0)\frac{\Delta^2}{4}\ . \tag{262}$$

Notice that we have halved the range of K integration, but doubled the number of variables at each K. The two-by-two matrix, which is traceless due to the relation

$$E(K')=-\cos(K+\pi)=-E(K)\ , \tag{263}$$

is readily diagonalized. The one-particle energy levels come in equal and opposite pairs, and we fill the negative energy states to obtain the following ground-state energy per unit volume:

$$\frac{E_0}{2\pi\delta(0)}=\frac{\Delta^2U_0}{4}-\int_0^\pi\frac{dK}{2\pi}\sqrt{E^2(K)+\Delta^2U_0^2}\ , \tag{264}$$

where the integral comes from the filled sea. Minimizing with respect to Δ, we obtain the relation

$$\Delta=\int_0^\pi\frac{dK}{\pi}\frac{U_0\Delta}{\sqrt{E^2(K)+\Delta^2U_0^2}}\ . \tag{265}$$

Assuming $\Delta\neq0$, we cancel it on both sides. It is clear that $U_0<0$ is not acceptable, since the two sides of the equation would then have opposite signs.

For positive U_0, a nontrivial solution requires that

$$1=U_0\int_0^\pi\frac{dK}{\pi}\frac{1}{\sqrt{E^2(K)+\Delta^2U_0^2}}\ ; \tag{266}$$

this is called the *gap equation.* On the left hand is the number 1, and on the right-hand side something of order U_0. It appears that we will get a solution only above some minimum U_0. This is wrong. The integrand becomes very large at the Fermi points $|K|=K_F=\pi/2$, where $E(K)$ vanishes. Writing

$$E(K)=k \ , \tag{267}$$

$$k=|K|-K_F \ , \tag{268}$$

we approximate the gap equation as follows:

$$1=U_0\int_{-\Lambda}^{\Lambda}\frac{dk}{\pi}\frac{1}{\sqrt{k^2+\Delta^2U_0^2}}\simeq\frac{2U_0}{\pi}\ln\frac{\Lambda}{\Delta U_0} \ , \tag{269}$$

where Λ, the upper cutoff on $|k|$, is not very important. What is important is that, due to the logarithmic behavior of the integral near the origin in k, i.e., near the Fermi surface, there will always be a solution to the gap equation given by

$$\Delta=\frac{\Lambda}{U_0}e^{-\pi/2U_0} \ . \tag{270}$$

The logarithmic divergence is also reflected in the divergent susceptibility of the noninteracting system to a probe (or perturbation) at momentum π. [At second order in perturbation theory, the perturbation will link the ground state to states of arbitrarily low energy in which a particle just below the right (left) Fermi point is pushed to just above the left (right) Fermi point. The small energy denominators, summed over such states, will produce the logarithm.]

Mean-field theory also predicts that the same thing will happen in $d=2$. In this case the nesting condition (readers unfamiliar with this concept should consult Fig. 17, Sec. X) ensures that the perturbation at (π,π) will excite particles just below the Fermi surface to states just above it on the "other side" *no matter where the starting point is on the Fermi surface.* Now for any Fermi surface, if we take some perturbation of any momentum (which is not too large compared to the size of the surface), there will always be some points just below sea level that will get scattered to points just above sea level. But these points would have to come from some special angular region of the Fermi surface which are connected by this momentum (the angular dimensions of this region will decrease with the energy denominators) and the integral over the small-energy denominators will converge. If we use a coordinate θ on the surface and a coordinate ε normal to it, the integrals will be of the form

$$\int_{-\Lambda}^{\Lambda}\int_{\theta_1(Q,\varepsilon)}^{\theta_2(Q,\varepsilon)}\frac{d\varepsilon\,d\theta}{\sqrt{\varepsilon^2+\Delta^2}} \ . \tag{271}$$

The cutoff Λ focuses on the small energy denominator region. In the absence of nesting, for a perturbation of some given momentum Q, the range of θ will shrink with ε and the integral will converge. On the other hand, at the nesting momentum, the angular integral will spread over the entire Fermi surface (since every point on or near the Fermi surface gets scattered to another, as shown in Fig. 17, Sec. X), and the integral will have a logarithmic divergence.

Returning to $d=1$, mean-field theory also predicts that the system will have a nonzero superconducting order parameter for the smallest attractive coupling. In the corresponding calculation the instability will stem from the time-reversal symmetry of the problem: $E(K)=E(-K)$.

Unfortunately both these predictions are wrong. The error comes from the neglected quartic operator $:n_j::n_{j+1}:$. We know all this because the present spinless Hamiltonian, Eq. (248), can be solved exactly (Yang and Yang, 1976).[10] The exact solution tells us that the system remains gapless for U_0 of either sign until it exceeds a minimum value of order unity. We shall now develop the RG approach to this problem and obtain results in harmony with this exact result.

B. The RG approach for $d=1$ spinless fermions

Our goal is to explore the stability of the noninteracting spinless fermions to weak interactions. We are interested not just in the fate of the model in Eq. (248) but in a whole family of models of which this will be special case. Our strategy, stated earlier on, is the following. First we argue that, at weak coupling, only modes near $\pm K_F$ will be activated. Thus we shall linearize the dispersion relation $E(K)=-\cos K$ near these points and work with a cutoff Λ,

$$H_0=\sum_i\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\psi_i^\dagger(k)\psi_i(k)k \ , \tag{272}$$

where

$$k=|K|-K_F \ , \tag{273}$$

$$i=L,R \ \text{(left or right)} \ . \tag{274}$$

Notice that H_0 is an integral over fermionic oscillators, which we studied in Sec. III. The frequency Ω_0 of the oscillator at momentum k is simply k.

Next we shall write down a $T=0$ partition function for the noninteracting fermions. This will be a Grassmann integral only over the degrees of freedom within a cutoff Λ of the Fermi surface. We shall then find an RG transformation that lowers the cutoff but leaves the free-field action, S_0, invariant. With the RG well defined, we shall look at the generic perturbations of this fixed point and classify them as usual. If no relevant operators show up, we shall still have a scale-invariant gapless system. If, on the other hand, there are generic relevant perturbations, we shall have to see to which new fixed point the system flows. (The new one could also be gapless.) The stability analysis can be done perturbatively. In particular, if a relevant perturbation takes us away from the original fixed point, *nothing at higher orders can ever bring us back to this fixed point.*

[10]The reader consulting this reference will find that these authors solve a problem of quantum spins on a line. This *XXZ chain* is related to the spinless fermions by a Jordan Wigner transformation.

Let us then begin with the partition function for our system of fermions:

$$Z_0=\int \prod_{i=L,R}\prod_{|k|<\Lambda} d\bar{\psi}_i(\omega k)d\psi_i(\omega k)e^{S_0}\ , \tag{275}$$

$$S_0=\sum_{i=L,R}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\bar{\psi}_i(\omega k)(i\omega-k)\psi_i(\omega k)\ . \tag{276}$$

This is just a product of functional integrals for the Fermi oscillators at each momentum with $\Omega_0(k)=k$. The first step in the RG transformation is to integrate out all $\psi(k\omega)$ and $\bar{\psi}(k\omega)$ with

$$\Lambda/s\le|k|\le\Lambda \tag{277}$$

and *all* ω. Thus our phase space has the shape of a rectangle, infinite in the ω direction, finite in the k direction. This shape will be preserved under the RG transformation. Since there is no real relativistic invariance here, we shall make no attempt to treat ω and k on an equal footing. Allowing ω to take all values allows us to extract an effective Hamiltonian operator at any stage in the RG, since locality in time is assured.

Since the integral is Gaussian, the result of integrating out fast modes is just a numerical prefactor, which we throw out. The surviving modes now have their momenta going from $-\Lambda/s$ to Λ/s. To make this action a fixed point, we define rescaled variables:

$$\begin{aligned} k'&=sk\ ,\\ \omega'&=s\omega\ ,\\ \psi_i'(k'\omega')&=s^{-3/2}\psi_i(k\omega)\ . \end{aligned} \tag{278}$$

Ignoring a constant that comes from rewriting the measure in terms of the new fields, we see that S_0 is invariant under the mode elimination and rescaling operations.

We can now consider the effect of perturbations on this fixed point. Rather than turn on the perturbation corresponding to the nearest-neighbor interaction we shall perform a more general analysis. The result for the particular case will be subsumed by this analysis.

C. Quadratic perturbations

First consider perturbations that are quadratic in the fields. These must necessarily be of the form

$$\delta S_2=\sum_{i=L,R}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\mu(k\omega)\bar{\psi}_i(\omega k)\psi_i(\omega k)\ , \tag{279}$$

assuming symmetry between left and right Fermi points.

Since this action separates into slow and fast pieces, the effect of mode elimination is simply to reduce Λ to Λ/s in the integral above. Rescaling moments and fields, we find

$$\mu'(\omega',k',i)=s\mu(\omega,k,i)\ . \tag{280}$$

We get this factor s as a result of combining a factor s^{-2} from rewriting the old momenta and frequencies in terms of the new and a factor s^3 which comes from rewriting the old fields in terms of the new.

Let us expand μ in a Taylor series,

$$\mu(k,\omega)=\mu_{00}+\mu_{10}k+\mu_{01}i\omega+\cdots+\mu_{nm}k^n(i\omega)^m+\cdots\ . \tag{281}$$

The constant piece is a relevant perturbation:

$$\mu_{00}\to s\mu_{00}\ . \tag{282}$$

This relevant flow reflects the readjustment of the Fermi sea to a change in chemical potential. The correct way to deal with this term is to include it in the free-field action by filling the Fermi sea to a point that takes μ_{00} into account. As for the next two terms, they clearly modify terms that are already present in the action, can be absorbed into them, and correspond to marginal interactions. When we consider quartic interactions, it will be seen that mode elimination will produce terms of the above form even if they were not there to begin with just as in ϕ^4 theory. The way to deal with them will be discussed in due course. As for higher-order terms in Eq. (281), they are irrelevant under the RG mentioned above. This is, however, a statement that is correct at the free-field fixed point. We shall have occasion to discuss a term that is irrelevant at weak coupling but gets promoted to relevance as the interaction strength grows.

D. Quartic perturbations: the RG at tree level

We now turn on the quartic interaction whose most general form is

$$\delta S_4=\frac{1}{2!2!}\int_{K\omega}\bar{\psi}(4)\bar{\psi}(3)\psi(2)\psi(1)u(4,3,2,1)\ , \tag{283}$$

where

$$\bar{\psi}(i)=\bar{\psi}(K_i,\omega_i),\ \text{etc.}, \tag{284}$$

$$\int_{K\omega}=\left[\prod_{i=1}^{4}\int_{-\pi}^{\pi}\frac{dK_i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_i}{2\pi}\right][2\pi\bar{\delta}(K_1+K_2-K_3-K_4)2\pi\delta(\omega_1+\omega_2-\omega_3-\omega_4)] \tag{285}$$

and $\bar{\delta}$ enforces momentum conservation mod2π, as is appropriate to any lattice problem. A process in which lattice momentum is violated in multiples of 2π is called an *umklapp* process. The delta function containing frequencies enforces time-translation invariance. The coupling function u is antisymmetric under the exchange of its first or last two arguments among themselves, since that is true of the Grassmann fields that it multiplies. *Thus*

the coupling u has all the symmetries of the full vertex function Γ with four external lines.

To get a feeling for all these ideas let us consider the nearest-neighbor repulsion from Eq. (248) and ask what u it generates in the action. Let us begin with the operator

$$H_I = U_0 \sum_j \psi_j^\dagger \psi_j \psi_{j+1}^\dagger \psi_{j+1} = -U_0 \sum_j \psi_j^\dagger \psi_{j+1}^\dagger \psi_j \psi_{j+1} \tag{286}$$

and make a Fourier transformation to get

$$H_I = -U_0 \left[\prod_{i=1}^{4} \int_{-\pi}^{\pi} \frac{dK_i}{2\pi} \right] 2\pi \bar{\delta}(K_1+K_2-K_3-K_4) \times \psi^\dagger(K_4)\psi^\dagger(K_3)\psi(K_2)\psi(K_1) e^{i(K_1-K_3)} . \tag{287}$$

We now antisymmetrize $e^{i(K_1-K_3)}$ with respect to $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$, since the rest of the integrand is antisymmetric under either of these operations. This gives us the result

$$e^{i(K_1-K_3)} \rightarrow \tfrac{1}{4}(e^{i(K_1-K_3)} - e^{i(K_2-K_3)} - e^{i(K_1-K_4)} + e^{i(K_2-K_4)}) . \tag{288}$$

We next use the fact that, due to the $\bar{\delta}$ function,

$$K_1+K_2 = K_3+K_4+Q = 0 \text{ or } 2\pi , \tag{289}$$

where Q is the umklapp momentum. It cannot be any higher multiple of 2π given that $|K_i| \le \pi$. This means that

$$\sin Q/2 = 0 . \tag{290}$$

This result and some simple trigonometry applied to Eq. (288) lead to the coupling

$$\frac{u(4,3,2,1)}{2!2!} = U_0 \sin\left[\frac{K_1-K_2}{2}\right] \sin\left[\frac{K_3-K_4}{2}\right] \times \cos\left[\frac{K_1+K_2-K_3-K_4}{2}\right] . \tag{291}$$

In arriving at this equation we have used Eq. (290) and gone from the interaction Hamiltonian to the corresponding action in the path integral in the usual way. The integral of $-H_I$ (with Fermi operators replaced by Grassmann numbers) from $\tau=0$ to $\tau=\infty$ becomes the integral over four frequencies with one overall delta function. Notice that u has no frequency dependence.

Let us now return to the general interaction, Eqs. (283)–(285), and restrict the momenta to lie within Λ of either Fermi point, L or R. Using a notation in which L (left Fermi point) and R (right Fermi point) become discrete, we find that a label $i = L$ or R and $1-4$ label the frequencies and momenta (measured from the appropriate Fermi points). Equations (283)–(285) become

$$\delta S_4 = \frac{1}{2!2!} \sum_{i_1 i_2 i_3 i_4 = L,R} \int_{K\omega}^{\Lambda} \bar{\psi}_{i_4}(4)\bar{\psi}_{i_3}(3)\psi_{i_2}(2)\psi_{i_1}(1) u_{i_4 i_3 i_2 i_1}(4,3,2,1) \tag{292}$$

where

$$\int_{K\omega}^{\Lambda} = \left[\int_{-\Lambda}^{\Lambda} \frac{dk_1 \cdots dk_4}{(2\pi)^4} \int_{-\infty}^{\infty} \frac{d\omega_1 \cdots d\omega_4}{(2\pi)^4} \right] [2\pi\delta(\omega_1+\omega_2-\omega_3-\omega_4)] \times [2\pi\bar{\delta}(\varepsilon_{i_1}(K_F+k_1)+\varepsilon_{i_2}(K_F+k_2)-\varepsilon_{i_3}(K_F+k_3)-\varepsilon_{i_4}(K_F+k_4))] \tag{293}$$

and

$$\varepsilon_i = \pm 1 \text{ for } R, L . \tag{294}$$

Let us now implement the RG transformation with this interaction. This proceeds exactly as in ϕ^4 theory. Let us recall how it goes. If schematically

$$Z = \int d\phi_< d\phi_> \exp[-\phi_<^2 - \phi_>^2] \exp[-u(\phi_< + \phi_>)^4] \tag{295}$$

is the partition function, and we are eliminating $\phi_>$, the effective u for $\phi_<$ has two origins. First, we have a term $-u\phi_<^4$, which is there to begin with, called the *tree-level* term. Next, there are terms generated by the $\phi_>$ integration. These are computed in a cumulant expansion and are given by Feynman diagrams whose internal momenta lie in the range being eliminated. The loops that contribute to the flow of u begin at order u^2.

Let us first do the order-u tree-level calculation for the renormalization of the quartic interaction. This gives us just Eq. (293) with $\Lambda \rightarrow \Lambda/s$. If we now rewrite this in terms of new momenta and fields, we get an interaction with the same kinematical limits as before, and we can meaningfully read off the coefficient of the quartic Fermi operators as the new coupling function. We find

$$u'_{i_4 i_3 i_2 i_1}(k_i', \omega_i') = u_{i_4 i_3 i_2 i_1}(k_i'/s, \omega_i'/s) . \tag{296}$$

The reader who carries out the intermediate manipulations will notice an important fact: K_F never enters any of the δ functions. Either all K_F's cancel in the nonumklapp cases, or they get swallowed up in multiples of 2π (in inverse lattice units) in the umklapp cases due to the periodicity of the $\bar{\delta}$ function. As a result the momentum δ functions are free of K_F and scale very nicely under the

RG transformation:

$$\bar{\delta}(k)\rightarrow\bar{\delta}(k'/s) \tag{297}$$

$$=s\bar{\delta}(k') \ . \tag{298}$$

Turning now to Eq. (296), if we expand u in a Taylor series in its arguments and compare coefficients, we find readily that the constant term u_0 is marginal and the higher coefficients are irrelevant. Thus u depends only on its discrete labels, and we can limit the problem to just a few coupling constants instead of the coupling function we started with. Furthermore, all reduce to just one coupling:

$$u_0=u_{LRLR}=u_{RLRL}=-u_{RLLR}=-u_{LRRL} \ . \tag{299}$$

Other couplings corresponding to $LL\rightarrow RR$ are wiped out by the Pauli principle, since they have no momentum dependence and cannot have the desired antisymmetry.

As a concrete example, consider the u that comes from the nearest-neighbor interaction Eq. (291), reproduced here for convenience,

$$\frac{u(4,3,2,1)}{2!2!}=U_0\sin\left[\frac{K_1-K_2}{2}\right]\sin\left[\frac{K_3-K_4}{2}\right]\times\cos\left[\frac{K_1+K_2-K_3-K_4}{2}\right] \tag{300}$$

and ask what sorts of couplings are contained in it.

If 1 and 2 are both from R, we find the following factor in the coupling:

$$\sin\left[\frac{K_1-K_2}{2}\right]=\sin\left[\frac{k_1-k_2}{2}\right]\simeq\left[\frac{k_1-k_2}{2}\right] , \tag{301}$$

which leads to the requisite antisymmetry but makes the coupling irrelevant (due to the k's). There will be one more power of k from 3 and 4, which must also come from near just one Fermi point so as to conserve momentum modulo 2π. For example, the unklapp process, in which $RR\leftrightarrow LL$, has a coupling

$$u_{NN}(\text{umklapp})\simeq(k_1-k_2)(k_3-k_4) \tag{302}$$

and is strongly irrelevant at the free-field fixed point.

On the other hand, if 1 and 2 come from opposite sides,

$$\sin\left[\frac{K_1-K_2}{2}\right]\simeq\sin[\pi/2+O(k)] \tag{303}$$

and likewise for 3 and 4, and we have a marginal interaction u_0 with no k's in the coupling.

The tree-level analysis readily extends to couplings with six or more fields. All these are irrelevant, even if we limit ourselves to constant (ω- and k-independent) couplings. To determine the ultimate fate of the coupling u_0, marginal at tree level, we must turn to the one-loop RG effects.

E. RG at one loop: The Luttinger liquid

Let us begin with the action with the quartic interaction and do a mode elimination. To order u, this leads to an induced quadratic term represented by the tadpole graph in Fig. 5. We set $\omega=k=0$ for the external legs and have chosen them to lie at L, the left Fermi point. The integral given by the diagram produces a momentum-independent term of the form $\delta\mu\bar{\psi}_L\psi_L$. But we began with no such term. Thus we do not have a fixed point in this case. Instead we must begin with some term $\delta\mu^*\bar{\psi}_L\psi_L$ such that, upon renormalization, it reproduces itself. We find it by demanding that

$$\delta\mu^*=s\left[\delta\mu^*-u_0^*\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{\Lambda/s<|k|<\Lambda}\frac{dk}{2\pi}e^{i\omega0^+}\frac{1}{i\omega-k}\right] , \tag{304}$$

where we have used the zeroth-order propagator and the fact that to this order any $u_0=u_0^*$. The exponential convergence factor is the one always introduced to get the right answer for, say, the ground-state particle density using $\langle\bar{\psi}\psi\rangle$. Doing the ω integral, we get

$$\delta\mu^*=s\left[\delta\mu^*-u_0^*\int_{\Lambda/s<|k|<\Lambda}\frac{dk}{2\pi}\theta(-k)\right] \tag{305}$$

$$=s\left[\delta\mu^*-\frac{\Lambda u_0^*}{2\pi}(1-1/s)\right] . \tag{306}$$

It is evident that the fixed point is given by

$$\delta\mu^*=\frac{\Lambda u_0^*}{2\pi} \ . \tag{307}$$

Alternatively, we could just as well begin with the following relation for the renormalized coupling:

$$\delta\mu'=s\left[\delta\mu-u_0^*\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{\Lambda/s<|k|<\Lambda}\frac{dk}{2\pi}e^{i\omega0^+}\frac{1}{i\omega-k}\right] , \tag{308}$$

which implies the flow

$$\frac{d\mu}{dt}=\mu-\frac{u_0^*}{2\pi} \ , \tag{309}$$

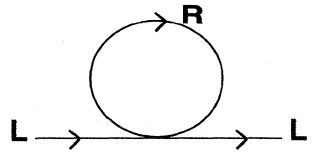

FIG. 5. The tadpole graph which renormalizes the fermion at one loop. It has no dependence on k (the deviation of the external momentum from K_F) or ω. We have used this freedom to set both these to zero on the external legs. The effect of this graph may be neutralized by a counterterm corresponding to a change in chemical potential. One may do this if one wants to preserve K_F.

assuming we choose to measure μ in units of Λ. The fixed point of this equation reproduces Eq. (307).

We can find $\delta\mu^*$ in yet another way with no reference to the RG. If we calculate the inverse propagator in the cutoff theory to order u, we shall find

$$G^{-1}=i\omega-k-\frac{\Lambda u_0}{2\pi}\ , \tag{310}$$

indicating that the Fermi point is no longer given by $k=0$. To reinstate the old K_F as interactions are turned on, we must move the chemical potential away from zero and to the value $\delta\mu=\Lambda u_0/2\pi$. Thus the correct action that gives us the desired K_F, for this coupling, to this order, is then schematically given by

$$S=\bar{\psi}(i\omega-k)\psi+\frac{\Lambda u_0}{2\pi}\bar{\psi}\psi+\frac{u_0}{2!2!}\overline{\psi\psi}\psi\psi\ . \tag{311}$$

An RG transformation on this action would not generate the tadpole graph contribution.

A very important point, which will appear again, is this: *we must fine-tune the chemical potential as a function of u, not to maintain criticality* (as one does in ϕ^4, where the bare mass is varied with the interaction to keep the system massless) *but to retain the same particle density.* [To be precise, we are keeping fixed K_F, the momentum at which the one-particle Green's function has its singularity. This amounts to keeping the density fixed (Luttinger, 1960).] If we kept μ at the old value of zero, the system would flow away from the fixed point with $K_F=\pi/2$, not to a state with a gap, but to another gapless one with a smaller value of K_F. This simply corresponds to the fact that, if the total energy of the lowest-energy particle that can be added to the system, namely μ, is to equal 0, the kinetic energy at the Fermi surface must be slightly negative so that the repulsive potential energy with the others in the sea brings the total to zero.

Now, we do not have to work with fixed density; we could take the given μ and accept whatever K_F it leads to. However, we shall frequently work at fixed density for two reasons. First, there is a simple experimental way to fix the density, namely, send in the desired number of particles and seal the system, i.e., work with the canonical and not the grand-canonical system. (This simple and viable procedure looks rather complicated in the grand-canonical language.) Secondly, we shall keep K_F fixed in many two- and three-dimensional cases to make contact with the pioneering work of Landau, done at fixed density.

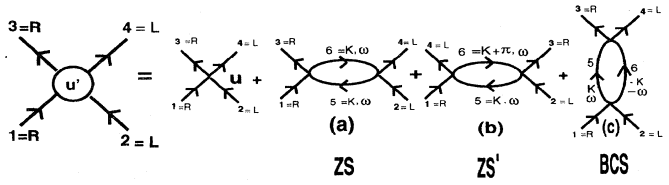

FIG. 6. The one-loop graphs for $\beta(u)$ for the spinless fermions in $d=1$. The loop momenta lie in the shell of width $d\Lambda$ being eliminated. The external frequencies being all zero, the loop frequencies are (a) equal for ZS, (b) equal for ZS', and (c) equal and opposite for the BCS graph. The ZS graph does not contribute, since both loop momenta are equal (the momentum transfer Q at the left vertex is 0) and lie a distance Λ above or below the Fermi surface, and the ω integration vanishes when the poles lie on the same half plane. The ZS' graph has momentum transfer π at the left and right ends. This changes the sign of the energy of the line entering at the left vertex. The ω integral is nonzero, the poles being on opposite half planes. The BCS graph (c) also survives, since the loop momenta are equal and opposite (because the incoming momentum is zero), and this again makes the poles go to the opposite half planes because the lines have *opposite* frequencies. The labels $1,\ldots,6$ refer to the master equation (312).

Let us now turn our attention to the order-u_0^2 graphs that renormalize u_0. These are shown in Fig. 6. The increment in u_0, hereafter simply called u, is given by the sum of the ZS (zero-sound), ZS', and BCS graphs. The analytical formula for the increment in u is

$$\begin{aligned}du(4321)=&\int u(6351)u(4526)G(5)G(6)\delta(3+6-1-5)d5\,d6-\int u(6451)u(3526)G(5)G(6)\delta(6+4-1-5)d5\,d6\\&-\frac{1}{2}\int u(6521)u(4365)G(5)G(6)\delta(5+6-1-2)d5\,d6\ ,\end{aligned} \tag{312}$$

where 1 to 4 stand for all the attributes of the (slow) external lines, 5 and 6 stand for all the attributes of the two (fast) internal lines—momenta (restricted to be within the region being eliminated) and frequencies—G are the propagators, the δ functions are for ensuring the conservation of momenta and frequencies, and $\int d5\,d6$ stands for sums and integrals over the attributes 5 and 6. The couplings u are functions of all these attributes, with all the requisite antisymmetry properties. (The order in which the legs are labeled in u is important, due to all the minus signs. The above equations have been written to hold with the indicated order of arguments. In their present form they are ready to be used by a reader who wants to include spin.)

This is the master formula, which we shall invoke often. It holds even in higher dimensions, if we suitably modify the integration region for the momenta.

To derive this formula, we do exactly what we did in Sec. II with bosons. We split the modes into slow and fast ones and do the fast integral, using the cumulant expansion to collect the terms that feed back to the quartic coupling. The calculation uses the fermionic Wick's theorem, and exactly the same three diagrams (ZS, ZS', and BCS) appear that we saw in the scalar field example.

The major difference from the scalar case is in the extra minus signs in the fermionic Wick's theorem. Of course the propagators now have different forms, and the range on loop variables is different. Readers familiar with Feynman diagrams may obtain this formula by drawing all the diagrams to this order in the usual Feynman graph expansion, but allowing the loop momenta to range only over the modes being eliminated. In the present case, these are given by the four heavy lines labeled a, b, c, and d in Fig. 7, where each line stands for a region of width $d\Lambda$ located at the cutoff, i.e., a distance Λ from the Fermi points. The external momenta are chosen to be $(4321)=(LRLR)$, at the Fermi surface. All the external k's and ω's are set equal to zero, since the marginal coupling u has no dependence on these. This has two consequences. First, the loop frequencies in the ZS and ZS' graphs are equal, while those in the BCS graph are equal and opposite. Second, the momentum transfers at the left vertex are $Q=K_1-K_3=0$ in the ZS graph and $Q'=K_1-K_4=\pi$ in the ZS' graph, while the total momentum in the BCS graph is $P=K_1+K_2=0$. Therefore, if one loop momentum $5=K$ lies in any of the four shells in Fig. 7, so does the other loop momentum 6, which equals K, $K+\pi$, or $-K$ in the ZS, ZS', and BCS graphs, respectively. Thus we may safely eliminate the momentum-conserving δ function in Eq. (312) using $\int d6$. This fact, coupled with

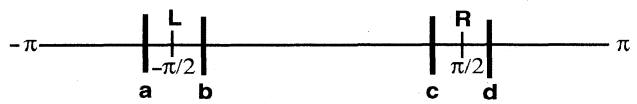

FIG. 7. Regions of momentum space being integrated out in the $d=1$ spinless fermion problem. The thick lines stand for slices of width $d\Lambda$. They lie a distance Λ from the Fermi points L and R. In the ZS graph, which has zero momentum transfer, both lines lie on the same slice and the ω integral gives zero. In the ZS' graph, the momentum transfer π connects a and c (which have opposite energies), and b and d are similarly related. In the BCS diagram the loop momenta are equal and opposite and correspond to a,d or b,c.

$$E(-K)=E(K)\ , \tag{313}$$

$$E(K'=K\pm\pi)=-E(K)\ , \tag{314}$$

leads to

$$du(LRLR)=\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\frac{u(KRKR)u(LKLK)}{[i\omega-E(K)][i\omega-E(K)]}-\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\frac{u(K'LKR)u(RKLK')}{[i\omega-E(K)][i\omega+E(K)]}$$
$$-\frac{1}{2}\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\frac{u(-KKLR)u(LR-KK)}{[i\omega-E(K)][-i\omega-E(K)]} \tag{315}$$

$$\equiv \mathrm{ZS}+\mathrm{ZS}'+\mathrm{BCS}\ , \tag{316}$$

where $\int_{d\Lambda}$ means the momentum must lie in one of the four slices in Fig. 7.

The reader is reminded once again that the names ZS, ZS', or BCS refer only to the topologies of the graphs. To underscore this point, especially for readers who have seen a similar integral in zero-sound calculations, we shall now discuss the ZS graph. In the present problem the loop momentum K lies within a sliver $d\Lambda$ of the cutoff. Both propagators have poles at the point $\omega=-iE(k=\pm\Lambda)$. No matter which half plane this lies in, we can close the contour the other way and the ω integral vanishes. This would be the case even if a small external momentum transfer ($Q=K_3-K_1\ll\Lambda$) took place at the left vertex, since both poles would still be on the same side. This is very different from what happens in zero-sound calculations, where the loop momenta roam freely within the cutoff and, in particular, go to the Fermi surface. In that case, the integral becomes very sensitive to how the external momentum transfer $Q=K_3-K_1$ and frequency transfer $\Omega=\omega_3-\omega_1$ are taken to zero, since any nonzero Q, however small, will split the poles and make them lie on different half planes for $k<Q$, and the integral will be nonzero. It is readily seen that

$$\int_{-\infty}^{\infty}\int_{-\Lambda}^{\Lambda}\frac{d\omega\,dk}{4\pi^2}\frac{1}{(i\omega-k)(i\omega-k-Q+i\Omega)}$$
$$=\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\frac{i}{\Omega+iq}[\theta(k)-\theta(k+q)]\ , \tag{317}$$

where the step function $\theta(k)$ is simply related to the Fermi function: $f(k)=1-\theta(k)$. If we keep $\Omega\neq 0$ and send Q to zero, we get zero. On the other hand, if we set $\Omega=0$ and let Q approach zero, we get (minus) the derivative of the (Fermi) θ function, i.e., a δ function at the Fermi surface. Thus readers used to zero-sound physics should not be disturbed by the fact that the ZS graph makes no contribution, since the connotation here is entirely different.[11]

[11]Before we move onto the ZS' graph, let us notice another related fact. Suppose we choose to find the one-loop β function using the field theory method. Then we calculate the four-point function Γ in a cutoff theory and demand that it be cutoff independent. The same three graphs will appear in the expression for Γ, but the loop integrals will in fact go all the way up to the cutoff. Consequently the ZS graph will make a contribution that is very sensitive to how the small external momenta and frequencies are chosen. However, since this contribution, when nonzero, comes from integrating a δ function at the Fermi surface, it will not have any sensitivity to the cutoff and will make no contribution to the derivative with respect to the cutoff, i.e., to the β function. The situation is parallel to what we saw in massless ϕ^4 theory in Sec. II. There the expression for Γ had an infrared divergence when $r_0=0$, but that did not affect the derivative with respect to the upper cutoff Λ. On the other hand, in the modern approach one never saw any singular behavior even in the intermediate steps.

Now for the ZS′ graph, Fig. 6(b), Eq. (315). We see that K must lie near L, since $1=R$ and there is no RR scattering. As far as the coupling at the left vertex is concerned, we may set $K=L$, since the marginal coupling has no k dependence. Thus $K+\pi=R$ and the vertex becomes $u(RLLR)=-u$. So does the coupling at the other vertex. Performing the ω integral (which is now nonzero, since the poles are always on opposite half planes), using the fact that there are two shells (a and b in Fig. 7) near L and that $|E(K)|=|k|=|\Lambda|$, we obtain

$$ZS'=u^2\int_{d\Lambda\in L}\frac{dK}{4\pi|E(K)|}=\frac{u^2}{2\pi}\frac{d|\Lambda|}{\Lambda}\ . \tag{318}$$

TABLE V. Spinless fermions in $d=1$: Summary of symbols and formulas.

CDW	Charge-density wave.
$H=-\frac{1}{2}\sum_j\psi^\dagger(j+1)\psi(j)+\text{H.c.}+U_0\sum_j[\psi^\dagger(j)\psi(j)-\frac{1}{2}][\psi^\dagger(j+1)\psi(j+1)-\frac{1}{2}]$	
$n_j=\psi_j^\dagger\psi_j$	Particle number at site j.
$\mu=U_0$	The chemical potential that ensures half filling for repulsion U_0.
$n_j=\frac{1}{2}+\frac{1}{2}(-1)^j\Delta+:n_j:$	The mean-field ansatz. H appears as follows in terms of Δ:
$H=-\frac{1}{2}\sum_j\psi^\dagger(j+1)\psi(j)+\text{H.c.}+U_0\left[\frac{1}{4}\sum_j\Delta^2-\Delta\sum_j(-1)^j n_j\right]+U_0\left[\sum_j:n_j::n_{j+1}:\right]$	
$:n_j:$	Charge density at site j with mean value subtracted.
$U_0\sum_j:n_j::n_{j+1}:$	Part neglected in mean-field calculation.
Δ	CDW order parameter.
$E(k)=-\cos K$	Dispersion relation for free fermions.
$K_F=\pi/2$	The Fermi momentum at half filling.
K	Momentum measured from origin.
$k=\lvert K\rvert-K_F$	Momentum measured relative to the Fermi surface (points).
L and R	Names for left and right Fermi points $K=\mp\pi/2$.
$H_0=-\int_{-\pi}^{\pi}\frac{dK}{2\pi}\psi^\dagger(K)\psi(K)\cos K$	Free-field Hamiltonian.
$H_0=\sum_i\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\psi_i^\dagger(k)\psi_k(k)k$	Free-field Hamiltonian linearized near K_F.
$Z_0=\int\prod_{i=L,R}\prod_{\lvert k\rvert<\Lambda}d\bar\psi_i(\omega k)d\psi_i(\omega k)e^{S_0}$	Free-fermion partition function.
$S_0=\sum_{i=L,R}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\bar\psi_i(\omega k)(i\omega-k)\psi_i(\omega k)$	Free-fermion action.
$k'=sk\quad\omega'=s\omega\quad\psi_i'(k'\omega')=s^{-3/2}\psi_i(k\omega)$	RG transformation.
$\delta S_2=\sum_i\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\mu(k\omega)\bar\psi_i(\omega k)\psi_i(\omega k)$	Quadratic perturbation.
$\mu(k,\omega)=\mu_{00}+\mu_{10}k+\mu_{01}i\omega+\cdots$	Taylor expansion of quadratic coupling.[a]
$\delta S_4=\frac{1}{2!2!}\int_{K\omega}\bar\psi(4)\bar\psi(3)\psi(2)\psi(1)u(4,3,2,1)$	Quartic perturbation in schematic form.
$\int_{K\omega}=\left[\prod_{i=1}^{4}\int_{-\pi}^{\pi}\frac{dK_i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_i}{2\pi}\right][2\pi\bar\delta(K_1+K_2-K_3-K_4)2\pi\delta(\omega_1+\omega_2-\omega_3-\omega_4)]$	
$\bar\delta$	Delta function for momentum conservation modulo 2π.
Umklapp process	A process in which momentum is changed by a multiple of 2π.
$u(4,3,2,1)=U_0\sin\left[\frac{K_1-K_2}{2}\right]\sin\left[\frac{K_3-K_4}{2}\right]\times\cos\left[\frac{K_1+K_2-K_3-K_4}{2}\right]$	Nearest-neighbor model.
$u'_{i_4i_3i_2i_1}(k_i',\omega_i')=u_{i_4i_3i_2i_1}(k_i'/s,\omega_i'/s)$	Renormalization of quartic coupling at tree level. $i=L,R$.
$u_0=u_{LRLR}=u_{RLRL}=-u_{RLLR}=-u_{LRRL}$	Symmetries of marginal coupling constant.
$\frac{du_0}{dt}=0$	Flow at one loop.

[a]Only μ_{00} is relevant. The rest are marginal or irrelevant.

The reader may wish to check that the ZS' graph will make the same contribution to the β function in the field theory approach.

The BCS graph [Eq. (315), Fig. 6(c)] gives a nonzero contribution, since the propagators have opposite frequencies and opposite momenta, but equal energies due to time-reversal invariance $E(K)=E(-K)$. We notice that the factor of $\frac{1}{2}$ is offset by the fact that K can now lie in any of the four regions a, b, c, or d. We obtain a contribution of the same magnitude as ZS' but opposite sign, so that

$$du=\left[\frac{u^2}{2\pi}-\frac{u^2}{2\pi}\right]\underbrace{\frac{d|\Lambda|}{\Lambda}}_{dt}, \tag{319}$$

$$\frac{du}{dt}=\beta(u)=0 . \tag{320}$$

Thus we find that u is *still* marginal. The flow to one loop for μ and u is

$$\frac{d\mu}{dt}=\mu-\frac{u}{2\pi}, \tag{321}$$

$$\frac{du}{dt}=0 . \tag{322}$$

There is a line of fixed points:

$$\mu=\frac{u^*}{2\pi}, \tag{323}$$

$$u^* \text{ arbitrary} . \tag{324}$$

Notice that β vanishes due to a cancellation between two diagrams, each of which by itself would have led to the CDW or BCS instability. When one does a mean-field calculation for CDW, one focuses on just the ZS' diagram and ignores the BCS diagram. This amounts to taking

$$\frac{du}{dt}=\frac{u^2}{2\pi}, \tag{325}$$

which, if correct, would imply that any positive u grows under renormalization. If this growth continues, we expect a CDW. On the other hand, if just the BCS diagram is kept, u would flow to large negative couplings leading to a state with $\langle\psi_R\psi_L\rangle\neq 0$.

What the β function does is treat these competing instabilities simultaneously and predict a scale-invariant theory.

Is this the correct prediction for the spinless model, which, as we saw, had the marginal interaction in its interaction? Yes, the exact solution of Yang and Yang (1976) tells us there is no gap until u is of order unity. If the RG analysis were extended to higher loops, we would keep getting $\beta=0$ to all orders. This follows from the Ward identity in the cutoff continuum model (Di Castro and Metzner, 1991; Metzner and Di Castro 1992), which reflects the fact that, in this model, the number of fermions of type L and R are separately conserved. How do we ever reproduce the eventual CDW instability known to exist in the exact solution? The answer is as follows. As we move along the line of fixed points, labeled by u, the dimensions of various operators will change from the free-field values. Ultimately the umklapp coupling ($RR\leftrightarrow LL$), which was suppressed by a factor $(k_1-k_2)(k_3-k_4)$, will become marginal and then relevant. If we were not at half filling, such a term would be ruled out by momentum conservation, and the scale-invariant state, called a Luttinger liquid (Luttinger, 1961; Haldane, 1981), would persist for all u. While this liquid provides us with an example of where the RG does better than mean-field theory, it is rather special and seems to occur in $d=1$ systems where the two Fermi points satisfy the conditions for *both* CDW and BCS instabilities. In higher dimensions we shall find that any instability due to a divergent susceptibility is not precisely canceled by another.

As an aside, note that in the ZS' and BCS diagrams, the integrand is a function of just ω^2+k^2, so that we have rotational (Euclidean) invariance. In this case we can, if we wish, work with a disk of radius Λ in ω-k space rather than the rectangle of width Λ and infinite height. You may check that if we integrate out a shell of thickness $d\Lambda$ in the ω-k space, we get the same contribution to the β function.

The main results from this section are summarized in Table V.

V. THE RG IN $d>1$: ROTATIONALLY INVARIANT CASE AT TREE LEVEL

We now proceed to apply exactly the same approach to spinless fermions in $d>1$. The nontrivial geometry of the Fermi surface will play a profound role, and the application of the RG leads to some phenomena not seen in the usual applications to critical phenomena, which is of course what makes it interesting. We start with the simplest case of a circular Fermi surface in $d=2$. The extension of the analysis to spherical surfaces in $d=3$ is very direct and will be explained. Only in Sec. X will we take up the nested Fermi surface that leads to CDW formation.

A. Tree level in $d=2$

Let us begin with a square lattice containing spinless fermions at very low filling. In this case we can approximate the free-particle dispersion relation as

$$E=-\cos K_x-\cos K_y \tag{326}$$

$$\simeq -2+K^2/2 . \tag{327}$$

Since the problem now has rotational invariance, it is isomorphic to the problem of electrons in free space with a dispersion relation

$$E=\frac{K^2}{2m} . \tag{328}$$

We shall therefore study the latter, *since this allows us to make contact with Landau's work on it.* Let us introduce a chemical potential so that the ground state is a Fermi circle of radius $K_F=\sqrt{2m\mu}$. Next, we linearize the dispersion relation near the Fermi surface:

$$\varepsilon(K)=E(K)-\mu \tag{329}$$

$$=\frac{K^2-K_F^2}{2m} \tag{330}$$

$$=\frac{kK_F}{m}+O(k^2) \quad (k=|K|-K_F) \tag{331}$$

$$\equiv vk \ , \tag{332}$$

where v is the Fermi velocity. The free-field action now becomes

$$S_0=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_0^{2\pi}\frac{d\theta}{2\pi}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}[\bar{\psi}(\omega\theta k)(i\omega-vk)\psi(\omega\theta k)] \ . \tag{333}$$

To obtain this, we must replace the measure $K\,dk$ by $K_F dk$, as the difference will prove irrelevant under the RG and absorb a factor $\sqrt{K_F}$ into each of the two Fermi fields. Mode elimination proceeds just as in $d=1$: we eliminate all modes obeying $\Lambda/s\le|k|\le\Lambda$ for all ω and θ. The same scaling of k and ω and the fields as in Eq. (278) leaves S_0 invariant. The only difference is that the internal index i which took just two values (L =left and R =right) is now replaced by a continuous parameter θ. *The $d=2$ theory thus looks like an integral over one-dimensional theories, one for each direction, each with infinitesimal weight.*

Next we dispense with rotationally invariant quadratic interactions as in $d=1$: either they modify the chemical potential, rescale existing terms, or are irrelevant. (Also irrelevant is the difference between K and K_F in the measure.) Let us now move on to the really interesting case of the quartic interaction. This has the general form

$$\delta S_4=\frac{1}{2!2!}\int_{K\omega\theta}\bar{\psi}(4)\bar{\psi}(3)\psi(2)\psi(1)u(4,3,2,1) \tag{334}$$

where

$$\bar{\psi}(i)=\bar{\psi}(K_i,\omega_i,\theta_i), \quad \text{etc.}, \tag{335}$$

$$\int_{K\omega\theta}=\left[\prod_{i=1}^{3}\int_0^{2\pi}\frac{d\theta_i}{2\pi}\int_{-\Lambda}^{\Lambda}\frac{dk_i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_i}{2\pi}\right]\theta(\Lambda-|k_4|) \ , \tag{336}$$

$$k_4=|\mathbf{K}_4|-K_F \ . \tag{337}$$

Much of the new physics stems from this measure for quartic interactions. We shall examine it in some detail, focusing on the factor $\theta(\Lambda-|k_4|)$, which plays a crucial role.

We start with a quartic interaction that is invariant under space-time translations and Fourier-transform it, getting an integral over four ω's subject to a δ-function constraint and four momenta $\mathbf{K}_i$ subject to a momentum-conserving δ function. Let us now eliminate one of the four sets of variables, say, the one numbered 4, by integrating them against the δ functions. The ω integral is easy: since all ω's are allowed, the condition $\omega_4=\omega_1+\omega_2-\omega_3$ is always satisfied for any choice of the first three frequencies. The same would be true for the momenta if all momenta were allowed. But they are not; they are required to lie within the annulus of thickness 2Λ around the Fermi circle. Consequently, if one freely chooses the first three momenta from the annulus, the fourth could have a length as large as $3K_F$. The role of $\theta(\Lambda-|k_4|)$ in Eq. (336) is to prevent exactly this.

Now such a θ function will arise in the ϕ^4 theory also if we eliminate $\mathbf{k}_4$ by integrating it against the momentum-conserving δ function. Its effect, however, is quite different there. For one thing, even if we ignore it, nothing very serious happens, since the first three k's lie in the tiny ball of size Λ and k_4 can never stray too far, being bounded by 3Λ. In particular, it will be controlled by Λ and decrease with it. In the present case, even if the first three momenta lie on the Fermi surface, the fourth can be off by an amount of order K_F rather than Λ. Secondly, even if we keep the θ function in the ϕ^4 case, its response to renormalization is very simple. Under the action of $\Lambda\to\Lambda/s$,

$$\theta(\Lambda-|\mathbf{k}_4|)\to\theta(\Lambda/s-|\mathbf{k}_4|) \tag{338}$$

$$=\theta(\Lambda-s|\mathbf{k}_4|) \tag{339}$$

$$=\theta(\Lambda-s|\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3|) \tag{340}$$

$$=\theta(\Lambda-|\mathbf{k}_4'|) \ . \tag{341}$$

Thus the θ function of the old variables goes into exactly the same function of the new variables. Since the rest of the integration measure goes into itself upon rescaling from k to k' (and absorbing factors of s into the new fields), we get the usual result,

$$u'(k')=u(k'/s) \ . \tag{342}$$

Upon doing a Taylor series in its arguments, we get the familiar result that the constant part u_0 is marginal and the other Taylor coefficients are irrelevant in $d=4$.

Let us try to do the same here, starting with Eqs. (334)–(337). We first reduce the range of each k_i by a factor s. Then we rescale all momenta to bring the range back to the old value. We must finally see if the θ function responds as it did above in the case of the ϕ^4 theory. If it does, we could conclude that

$$u'(k',\omega',\theta)=u(k'/s,\omega'/s,\theta) \ .$$

But it does not! The problem is that k_4 is a function not just of the other three little k's but also of K_F:

$$k_4=|(K_F+k_1)\mathbf{\Omega}_1+(K_F+k_2)\mathbf{\Omega}_2-(K_F+k_3)\mathbf{\Omega}_3|-K_F \ , \tag{343}$$

where $\mathbf{\Omega}_i$ is a unit vector in the direction of $\mathbf{K}_i$:

$$\mathbf{\Omega}_i = \mathbf{i}\cos\theta_i + \mathbf{j}\sin\theta_i \ , \tag{344}$$

where θ_i is the orientation of the unit vector along momentum $\mathbf{K}_i$. (In $d=2$, we shall use $\mathbf{\Omega}_i$ and θ_i interchangeably.)

It is now easy to check that if we carry out the manipulation that led to Eq. (341) we shall find

$$\theta(\Lambda - |k_4(k_1,k_2,k_3,K_F)|) \rightarrow \theta(\Lambda - |k_4'(k_1',k_2',k_3',sK_F)|) \ . \tag{345}$$

Thus the θ function after the RG transformation is not the same function of the new variables as the θ function before the RG transformation was of the old variables, due to the fact that $K_F \rightarrow sK_F$. As mentioned earlier, we cannot ignore the θ function, since in contrast to ϕ^4 theory, it is possible for very large k_4 (of order K_F) to arise, even if the first three are of order Λ. We have a real problem implementing the RG program: how are we to say what the new coupling is if the integration measure does not come back to its old form?

Before describing the solution to this impasse, let us restate the problem in more geometric terms. Imagine that we have renormalized with a large s and are down to a shell of very small thickness, i.e., Λ/K_F is very tiny. Thus all the momenta are essentially on the Fermi surface, and the only freedom is in their direction, $\mathbf{\Omega}_i$, or θ_i. *Now the point is that we cannot choose three of these angles freely, but only two, if all vectors are to lie on the Fermi circle.* For example, if we choose angles θ_1 and θ_2, the sum of the corresponding vectors lies along the bisector of the angle between them. The only way this initial-state momentum can equal the final-state momentum $\mathbf{K}_3+\mathbf{K}_4$ is for the final angles to equal the initial angles:

$$\text{Case I:} \quad \theta_3 = \theta_1 \ , \tag{346}$$

$$\theta_4 = \theta_2 \tag{347}$$

or

$$\text{Case II:} \quad \theta_3 = \theta_2 \ , \tag{348}$$

$$\theta_4 = \theta_1 \ . \tag{349}$$

In the case of identical spinless fermions, Cases I and II are physically equivalent.

There is only one exception: if the initial angles are exactly opposed to each other, leading to a total momentum $\mathbf{P}=0$, the final momenta are free to point in any direction as long as they oppose each other:

$$\text{Case III:} \quad \theta_2 = -\theta_1 \ , \tag{350}$$

$$\theta_4 = -\theta_3 \ . \tag{351}$$

In summary, *either 3 and 4 are slaved to (required to be equal to) 1 and 2* or *2 and 4 are slaved to (required to be opposite to) 1 and 3.*

Let us now back off from the limit $\Lambda/K_F=0$ and discuss the problem when Λ/K_F is small but not zero. Figure 8 depicts the situation. First let us ask for all pairs of momenta that lie within the annulus, and add up to some $\mathbf{P}$. The construction in Fig. 8 gives all of them. First we draw two annuli with centers separated by $\mathbf{P}$. They intersect in two regions (called I and II in the figure) of size of order Λ in each direction. If we start at the center of the left annulus and draw a vector to any point in I or II, and then a vector from that point to the center of the right annulus, we get two vectors that meet the twin requirements listed above. For example, the initial vectors $\mathbf{K}_1$ and $\mathbf{K}_2$ correspond to choosing this point from region II. Since $\mathbf{K}_3+\mathbf{K}_4=\mathbf{P}$, the latter pair must also stem from this construction. The figure shows them linked to region I. It is clear that the direction of the final vectors is within Λ/K_F of the initial vectors, with $\theta_3 \simeq \theta_2$ and $\theta_4 \simeq \theta_1$ for this choice. (Had we chosen the final vectors to come from region II also, the other pairing would have occurred.) The figure makes it clear that if we choose the orientations of the first three vectors without paying attention to this restriction, i.e., choosing a point outside regions I and II but inside the annulus for terminating $\mathbf{K}_3$, $\mathbf{K}_4$ can end up being much shorter or longer than K_F.

The problem of the changing θ function may be stated as follows in terms of this figure. Let us begin with some cutoff and a choice of four angles that is allowed at that cutoff as per this construction. There is some coupling u for this choice of angles. If we reduce the cutoff, the allowed choices of angles shrinks, and a coupling that was previously allowed may no longer be allowed. Since the angles θ play the role of internal (isospin-like) degrees of freedom, we have a situation in which the range of internal labels is changing under mode elimination. This makes it impossible to make a meaningful comparison of the couplings before and after, since certain processes are no longer allowed. [Imagine an SU(4) theory renormaliz-

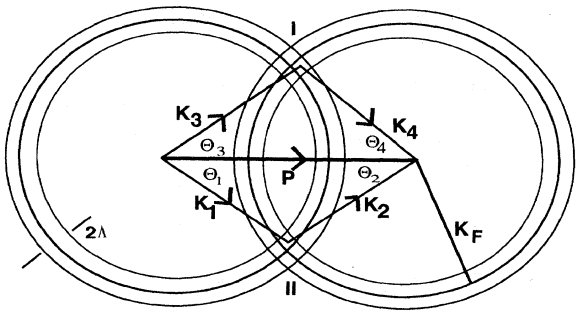

FIG. 8. The geometric construction for determining the allowed values of momenta. If K_1 and K_2 add up to P, then K_3 and K_4 are constrained as shown, if they are to add up to P and lie within the cutoff. Note that both the magnitude of K_3 *and its direction* θ_3 are constrained to lie within $\simeq \Lambda/K_F$ of θ_1. Had we chosen a K_3 that terminated in intersection II, θ_3 would have been within Λ/K_F of θ_2. At the fixed point, the angles become equal pairwise. If the incoming momenta K_1 and K_2 are equal and opposite, the two shells coalesce and K_3 and K_4 are free to point in all directions, as long as they are equal and opposite.

ing to an SU(3) theory.]

Since we do not want the momentum or internal labels to change their allowed range of values (if we are going to follow the standard RG procedure of comparing the couplings before and after to see how they are flowing), *we take the view that all the θ's are allowed, but that some u's abruptly renormalize to zero as the cutoff is reduced.*

The situation is a lot clearer if we use a smooth cutoff for k_4,

$$\theta(\Lambda-|k_4|)\to e^{-|k_4|/\Lambda}\ , \tag{352}$$

for now the process is not disallowed, but only exponentially suppressed as we renormalize. We may interpret that as the exponential decay of the corresponding u under renormalization.[12]

Now recall

$$k_4=|K_F(\underbrace{\mathbf{\Omega}_1+\mathbf{\Omega}_2-\mathbf{\Omega}_3}_{\Delta})+k_1\mathbf{\Omega}_1+k_2\mathbf{\Omega}_2-k_3\mathbf{\Omega}_3|-K_F\ , \tag{353}$$

where $\mathbf{\Omega}_i$ is a unit vector in the direction of $\mathbf{K}_i$. In what follows we shall keep just the Δ piece and ignore the $O(k)$ terms. This is because the only time the latter are comparable to the former is when both are of order Λ, in which case $k_4\simeq K_F$, and this region is exponentially suppressed by the smooth cutoff Eq. (352) anyway.

Under the RG transformation at tree level,

$$\prod_{i=1}^{3}\int_{-\Lambda}^{\Lambda}\frac{dk_i}{2\pi}\int_0^{2\pi}\frac{d\theta_i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_i}{2\pi}e^{-(K_F/\Lambda)\|\Delta|-1|}u(k\omega\theta)\bar{\psi}\psi\bar{\psi}\psi$$
$$\to\prod_{1}^{3}\int_{-\Lambda}^{\Lambda}\frac{dk_i'}{2\pi}\int_0^{2\pi}\frac{d\theta_i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_i'}{2\pi}e^{-(sK_F/\Lambda)\|\Delta|-1|}u\left[\frac{k'}{s}\frac{\omega'}{s}\theta\right]\bar{\psi}\psi\bar{\psi}\psi\ . \tag{354}$$

Let us write

$$e^{-(SK_F/\Lambda)\|\Delta|-1|}=e^{-(K_F/\Lambda)\|\Delta|-1|}e^{-[(s-1)K_F/\Lambda]\|\Delta|-1|}\ , \tag{355}$$

so that the measures before and after have the same factor $e^{-(K_F/\Lambda)\|\Delta|-1|}$. Now that the measures are identical before and after the RG transformation, we can compare apples to apples and identify the new quartic coupling:

$$u'(k'\omega'\theta)=e^{-[(s-1)K_F/\Lambda]\|\Delta|-1|}u\left[\frac{k'}{s}\frac{\omega'}{s}\theta\right]\ . \tag{356}$$

We may conclude that the only couplings that survive the RG transformation without any decay correspond to the cases in which

$$|\Delta|=|\mathbf{\Omega}_1+\mathbf{\Omega}_2-\mathbf{\Omega}_3|=1\ . \tag{357}$$

In $d=2$ this equation has only three solutions:

Case I: $\mathbf{\Omega}_3=\mathbf{\Omega}_1$ (hence $\mathbf{\Omega}_2=\mathbf{\Omega}_4$) , (358)

Case II: $\mathbf{\Omega}_3=\mathbf{\Omega}_2$ (hence $\mathbf{\Omega}_1=\mathbf{\Omega}_4$) , (359)

Case III: $\mathbf{\Omega}_1=-\mathbf{\Omega}_2$ (hence $\mathbf{\Omega}_3=-\mathbf{\Omega}_4$) . (360)

This result, which was anticipated earlier, can be deduced algebraically or seen geometrically by considering Fig. 8 in the limit of zero shell thickness, since this gives the allowed region in the hard cutoff scheme after an infinite amount of renormalization, i.e., at the fixed point.

Notice that we have an extra conservation law at the fixed point $\Lambda/K_F=0$. Not only is total momentum conserved, the set of individual momenta is also conserved. The only exception is when the incoming momenta are equal and opposite. Then they add up to zero, and the final momenta can be any two opposite points on the Fermi circle.

For couplings that do obey this condition, Eq. (356) becomes

$$u'(k'\omega'\theta)_{|\Delta|=1}=u(k'/s\omega'/s\theta)_{|\Delta|=1}\ . \tag{361}$$

Performing a Taylor expansion[13] in k and ω and comparing coefficients of separate powers, we conclude that the leading term, with no dependence on either variable, is marginal, while all the rest are irrelevant. We shall refer to this term as u hereafter.

We see that the tree-level fixed point is characterized by two independent functions and not a handful of couplings. They are

[12]Nothing is gained by using a soft cutoff for the rest.

[13]A common concern frequently expressed by those familiar with the many-body physics of these fermions is this: is there not some ambiguity as to how the limit $\omega\to 0$ is to be taken as $k_i\to 0$? If so, how does u even have a Taylor expansion at the origin of ωk space? The answer is that u is just the bare coupling that goes into the cutoff theory. It is not the full four-point function Γ which is calculated in the cutoff theory by summing all Feynman diagrams with u as the coupling, Λ as the cutoff, and propagators whose momenta can go right up the Fermi space. Indeed, the Γ so computed has all the above-mentioned singularities. In contrast, u is obtained by taking some analytic (short-range) interaction in the full momentum space and eliminating the modes *outside* the cutoff. This procedure cannot produce any nonanalyticity. The situation is like that in ϕ^4 theory, where the bare couplings in the action are analytic functions of momenta while the full Green's functions are plagued with infrared singularities coming from the soft modes.

$$u[\theta_4(\theta_1\theta_2\theta_3)=\theta_2;\theta_3=\theta_1;\theta_2;\theta_1]=F(\theta_1;\theta_2)=F(\theta_1-\theta_2\equiv\theta_{12}) \text{ (rot inv)}, \tag{362}$$

$$u[\theta_4(\theta_1\theta_2\theta_3)=\theta_1;\theta_3=\theta_2;\theta_2;\theta_1]=-F(\theta_{12}) \text{ (Pauli)}, \tag{363}$$

$$u[\theta_4(\theta_1\theta_2\theta_3)=-\theta_3;\theta_3;-\theta_1;\theta_1;]=V(\theta_1;\theta_3)=V(\theta_1-\theta_3\equiv\theta_{13}) \text{ (rot inv)}. \tag{364}$$

Note that the manifestation of the Pauli principle on F and V is somewhat subtle: F will not be antisymmetric under $1\leftrightarrow 2$ since, according to the way it is defined above, we cannot exchange 1 and 2 without exchanging 3 and 4 at the same time. On the other hand, since 3 and 4 can be exchanged without touching 1 and 2 in the definition of V, V must go to $-V$ when $\theta_{13}\to\theta_{13}+\pi$.

A concrete example is useful here. Let us begin with the nearest-neighbor interaction of spinless fermions in $d=2$, transcribed into momentum space,

$$\frac{u(\mathbf{K}_4,\ldots,\mathbf{K}_1)}{2!2!}=U_0\left[\sin\left[\frac{K_{1x}-K_{2x}}{2}\right]\sin\left[\frac{K_{3x}-K_{4x}}{2}\right]\cos\left[\frac{K_{1x}+K_{2x}-K_{3x}-K_{4x}}{2}\right]+x\leftrightarrow y\right], \tag{365}$$

and expand it to second order in K (as we did in the kinetic energy), obtaining a rotationally invariant interaction:

$$u\simeq(\mathbf{K}_1-\mathbf{K}_2)\cdot(\mathbf{K}_3-\mathbf{K}_4). \tag{366}$$

If we evaluate this with all $|K|$ at K_F and use Eqs. (362), (363), and (364), we obtain

$$F(\theta_{12})\simeq U_0|\mathbf{K}_1-\mathbf{K}_2|^2\simeq U_0(1-\cos\theta_{12}), \tag{367}$$

$$V(\theta_{13})\simeq U_0\mathbf{K}_1\cdot\mathbf{K}_3\simeq U_0\cos\theta_{13}. \tag{368}$$

Observe the effect of particle exchange on these functions.

Let us understand the significance of the coupling function F. If we calculate the four-point vertex Γ, it will be given to lowest order by u. Since the only u's that survive require that the final directions θ_3 and θ_4 equal the initial ones θ_1 and θ_2 up to a permutation, only forward scattering exists at the fixed point.[14] The only exception occurs when the incoming particles have opposite directions: then they can scatter to another pair with equal and opposite momenta, the corresponding amplitude being given, of course, by V.

When we study noncircular Fermi surfaces, we shall find yet another coupling function W that survives if there is nesting. This will correspond to processes in which the momentum transfer equals the nesting vector $\mathbf{Q}_N$.

B. Tree-level analysis in $d=3$

Let us repeat the preceding analysis in $d=3$, where the Fermi surface is parametrized by two angles θ and ϕ. All integrals over θ get replaced by integrals over a solid angle. The process of mode elimination and rescaling of fields and momenta k proceeds exactly as in $d=1$ or $d=2$, since the coordinates on the Fermi surface play the role of internal variables (like isospin) and are unaffected by the RG transformation. We then end up with Eq. (356). The equation for the couplings that survive is still the same as Eq. (357), but now the unit vectors $\mathbf{\Omega}_i$ can point anywhere in three dimensions. This in turn means that the condition $|\Delta|=1$ has not only the solutions given in Eqs. (358), (359), and (360) but a continuum of others. First consider Eqs. (358) and (359), which tell us that $\mathbf{\Omega}_3$ and $\mathbf{\Omega}_4$ must coincide with $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ up to a permutation. In $d=3$, the former can rotate about their sum. In other words, Fig. 8 with zero shell thickness must now be viewed as depicting, not the intersection of two circles, but two spheres. Thus the vectors 3 and 4 do not have to coincide with 1 and 2 (up to a permutation), but can rotate around a cone with opening angle equal to that between 1 and 2. The planes containing 1 and 2 can have an angle $\phi_{12,34}$ with the plane containing 3 and 4.

Let us review this. The incoming particle momenta 1 and 2 lie on the Fermi sphere. Their sum lies in the plane they define and bisects the angle between them. The final particles, also on the sphere, can give the same sum by lying anywhere on the cone generated by rotating the incoming pair around their sum.

Although the individual momenta are no longer conserved, we have the additional requirement that *the angle between the final pair be the same as the angle between the initial pair:*

$$\mathbf{\Omega}_1\cdot\mathbf{\Omega}_2=\mathbf{\Omega}_3\cdot\mathbf{\Omega}_4. \tag{369}$$

Once again the only exception is when the incoming momenta add up to zero. In this case the final momenta are free to point in any direction, as long as they are mutually opposite. *Thus couplings corresponding to nonforward scattering (in which the initial and final directions are not the same) do not vanish under the RG transformation, but survive as marginal couplings.* The fixed point is then characterized by a function

$$F=F(z_{12},\phi_{12;34}) \tag{370}$$

$$\equiv F(z,\phi) \tag{371}$$

of two variables,

$$z_{12}=\mathbf{\Omega}_1\cdot\mathbf{\Omega}_2$$

[14] This is true at lowest order when $\Gamma=u$. But higher orders in the forward-scattering coupling can only give forward scattering.

TABLE VI. Phase space for various fixed points: Summary of symbols and formulas. In all cases the RG transformation reduces Λ (limit on k or ε) by a factor $s=e^{-t}$.

One-particle phase space
$d=1$: The Fermi surface is a pair of points, called L (left) and R (right). phase space $=\sum_{i=L}^{R}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}$, $k=\lvert K\rvert-K_F$.
$d=2$: Circular case. The Fermi surface is parametrized by angle θ. phase space $=\int_0^{2\pi}\frac{d\theta}{2\pi}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}$, $k=\lvert K\rvert-K_F$.
$d=2$: Noncircular case. The Fermi surface is parametrized by an angle θ if connected and by an additional label $\alpha=\pm1$ if it has two branches. phase space $=\sum_{\alpha}\int_0^{2\pi}\frac{d\theta}{2\pi}J(\theta)\int_{-\Lambda}^{\Lambda}\frac{d\varepsilon}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}$, where J is the Jacobian on the Fermi surface and ε measures the energy relative to the Fermi surface.
$d=3$: Spherical case. The Fermi surface is parametrized by $z=\cos\theta$ and ϕ. phase space $=\int\frac{d\Omega}{4\pi^2}\int_{-\Lambda}^{\Lambda}\frac{dk}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}$ where $d\Omega=d\phi\,dz=d\phi\,d\cos\theta$.
The quartic interaction is written in schematic form as[a] $\delta S_4=\int_\Lambda\bar\psi(4)\bar\psi(3)\psi(2)\psi(1)u(4321)$

[a]The labels 1 and 4 stand for all the attributes of the fields: momentum, frequency, and spin if included. The coupling obeys $u(4321)=-u(3421)=-u(4312)=u(3412)$ due to Fermi statistics. $\int_\Lambda$ stands for the integral over the above phase space for each of the four fields, restricted by delta functions that impose ω and momentum conservation (modulo 2π if appropriate).

and $\phi_{12;34}$, which is the angle between the planes containing 12 and 34, respectively.

In addition to this we still have the V function coming from Eq. (360) with $V=V(\mathbf{\Omega}_1\cdot\mathbf{\Omega}_3)$.

If we go back to nearest-neighbor coupling evaluated in this case we find

$$F=U_0(1-z_{12})\cos\phi\ , \tag{372}$$

$$V=U_0 z_{13}\ . \tag{373}$$

Notice that, since 3 and 4 are not slaved to 1 and 2 anymore, we can exchange just the latter. This causes $\phi\to\phi+\pi$, which in turn changes the sign of F. As for V, it changes sign under $z_{13}\to-z_{13}$.

Tables VI and VII contain a list of fixed points, couplings, and their flows, including those to be discussed later in the paper.

VI. RG IN $d>1$: ROTATIONALLY INVARIANT CASE AT ONE LOOP

We have found that the couplings F and V are marginal at tree level. The next thing to do is to buckle down, as we did in $d=1$, and go to the one-loop graphs and see if they tilt the marginal ones towards relevance or irrelevance or preserve marginality. We first do the analysis in $d=2$ and then discuss briefly the changes encountered upon going to $d=3$.

A. RG for the two-point function

Just as in $d=1$, mode elimination produces at one loop (tadpole graph as in Fig. 5) a nonzero change in the quadratic term in the form of a chemical potential. To retain the old Fermi surface, we must find an input $\delta\mu^*$ that will reproduce itself under the RG transformation. Carrying out the same analysis as in $d=1$, we find

$$\delta\mu^*(\omega k\theta)=\int\frac{d\omega' dk' d\theta'}{(2\pi)^3}\frac{F(\theta-\theta')}{i\omega-v^*k} \tag{374}$$

$$=\int_{-\Lambda}^{0}\frac{dk'}{2\pi}\int_0^{2\pi}\frac{d\theta'}{2\pi}F(\theta-\theta') \tag{375}$$

$$=\delta\mu^*\ , \tag{376}$$

where the last equation signifies that $\delta\mu^*$ is a constant independent of ω, k, or θ.

We may also see this as the counterterm one must add as the interaction is turned on, to maintain the same Fermi surface. The same holds in $d=3$:

$$\delta\mu^*(\omega k\mathbf{\Omega})=\int_{-\Lambda}^{0}\frac{dk'}{2\pi}\int\frac{d\Omega'}{4\pi^2}F(z=\mathbf{\Omega}'\cdot\mathbf{\Omega},\phi=0) \tag{377}$$

$$=\delta\mu^*\ . \tag{378}$$

Note that only the forward-scattering F (with $\phi=0$) enters this equation.

B. The one-loop β function for F

Now we turn to the real issue: the renormalization of the quartic couplings. The analysis will be done using the modern approach, though a passing remark may occasionally be made about the field theory approach. Equation (312) gives in this case (suppressing vector symbols)

$$du(4321)=\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}\frac{u(K+Q,3,K,1)u(4,K,2,K+Q)}{[i\omega-E(K)][i\omega-E(K+Q)]}$$
$$-\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}\frac{u(K+Q',4,K,1)u(3,K,2,K+Q')}{[i\omega-E(K)][i\omega-E(K+Q')]}$$
$$-\frac{1}{2}\int_{-\infty}^{\infty}\int_{d\Lambda}\frac{d\omega\,dK}{4\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}\frac{u(P-K,K,2,1)u(4,3,P-K,K)}{[i\omega-E(K)][-i\omega-E(P-K)]} \tag{379}$$
$$\equiv \mathrm{ZS}+\mathrm{ZS}'+\mathrm{BCS}\,, \tag{380}$$

where $Q=K_3-K_1$, $Q'=K_4-K_1$, and $P=K_1+K_2$. The subscript $d\Lambda$ on $\int_{d\Lambda}$ reminds us that both loop momenta must be in the shells being eliminated. Although K is restricted to this region explicitly, the restriction on the other momentum—$K+Q$ in the ZS graph; $K+Q'$ in the ZS′ graph, and $P-K$ in the BCS graph—is implicit. The Feynman diagrams are shown in Fig. 9.

On the left-hand side of the above equation, we choose the variables 4321 such that the corresponding coupling survives renormalization at tree level, which means it must be an F or a V.

Let us begin with the renormalization of F. We set all

TABLE VII. Fixed-point couplings and flows: a summary. The marginal part of coupling u depends only on the angles that parametrize the Fermi surface. Subscript NN denotes values in nearest-neighbor model.

Fermi surface	Couplings marginal at tree level and their flow at one loop
A: $d=2$: Circular Fermi surface	$u(\theta_4=\theta_2,\theta_3=\theta_1,\theta_2,\theta_1)=-u(\theta_4=\theta_1,\theta_3=\theta_2,\theta_2,\theta_1)=F(\theta_1,\theta_2)=F(\theta_1-\theta_2)\equiv F(\theta_{12})$ $F_{NN}=U_0(1-\cos\theta_{12})$ $dF/dt=0.$ $u(\theta_4=-\theta_3,\theta_3,\theta_2=-\theta_1,\theta_1)=V(\theta_3-\theta_1)\equiv V(\theta_{13})$ $\frac{dV(\theta_l-\theta_3)}{dt}=-\frac{1}{8\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}V(\theta_1-\theta)V(\theta-\theta_3)$ $V_{NN}=U_0\cos\theta_{13}$ $\frac{dV_l}{dt}=-\frac{V_l^2}{4\pi}, V_l=\int_0^{2\pi}\frac{d\theta}{2\pi}e^{il\theta}V(\theta)$. (BCS instability if $V_l<0.$)
B: $d=2$: Noncircular Time-reversal invariant Assumed connected	F same as in A, but $F(\theta_1,\theta_2)\neq F(\theta_1-\theta_2)$. $dF/dt=0.$ V same as in A but $V(\theta_3,\theta_1)\neq V(\theta_3-\theta_1)$ $\frac{dV(\theta_1;\theta_3)}{dt}=-\frac{1}{8\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}V(\theta_1;\theta)V(\theta;\theta_3)J(\theta)$
C: $d=2$: Noncircular No time-reversal invariance	F is as in B. $dF/dt=0.$ No V. No BCS instability.
D: $d=2$: nested α labels branch r: hopping anisotropy	F defined as in B and $dF/dt=0.$ $F_{NN}=U_0\left[\sin^2\left(\frac{\theta_1-\theta_2}{2}\right)+\frac{1}{2}\left(1-\frac{\cos\theta_1\cos\theta_2}{r^2}-\frac{\alpha_1\alpha_2}{r^2}\sqrt{r^2-\cos^2\theta_1}\sqrt{r^2-\cos^2\theta_2}\right)\right]$ V defined as in B. Same flow as in B. $V_{NN}(\theta_3\alpha_3\theta_1\alpha_1]=U_0\left[\sin\theta_1\sin\theta_3+\frac{\alpha_1\alpha_3}{r^2}\sqrt{r^2-\cos^2\theta_1}\sqrt{r^2-\cos^2\theta_3}\right]$ $W[\theta_2\alpha_2;\theta_1\alpha_1]=u[\theta_2+\pi,-\alpha_2;\theta_1+\pi,-\alpha_1;\theta_2\alpha_2;\theta_1\alpha_1]\equiv u(\theta_2',\alpha_2',\theta_1'\alpha_1';\theta_2\alpha_2,\theta_1\alpha_1)$ $\frac{dW[\alpha_2\theta_2\alpha_1\theta_1]}{dt}=-\int\sum_{\alpha}W[\alpha_2\theta_2\alpha'\theta']W[\theta\alpha\theta_1\alpha_1]\frac{J(\theta)d\theta}{(2\pi)^2}$ $W_{NN}=-U_0\left[\sin^2\left(\frac{\theta_1-\theta_2}{2}\right)+\frac{1}{2}\left(1-\frac{\cos\theta_1\cos\theta_2}{r^2}-\frac{\alpha_1\alpha_2}{r^2}\sqrt{r^2-\cos^2\theta_1}\sqrt{r^2-\cos^2\theta_2}\right)\right]$
E: $d=3$: Spherical Ω_i is direction on sphere. $\phi_{12;34}$ is angle between 1–2 and 3–4 planes	$u(\Omega_4,\Omega_3,\Omega_2,\Omega_1)\vert_{\Omega_2\cdot\Omega_1=\Omega_4\cdot\Omega_3}=F(\Omega_1\cdot\Omega_2;\phi_{12;34})\equiv F(z,\phi)$. $dF/dt=0.$ $F(z,0)$ is Landau's F function. $F_{NN}=U_0(1-z)\cos\phi.$ $u(-\Omega_3,\Omega_3,-\Omega_1,\Omega_1)=V(\Omega_3\cdot\Omega_1)\equiv V(z_{13})$ $V_{NN}=U_0z_{13}=U_0\Omega_1\cdot\Omega_3$ $\frac{dV_l}{dt}=-cV_l^2$, $c>0$ and $V_l=\int V(z)P_l(z)dz$

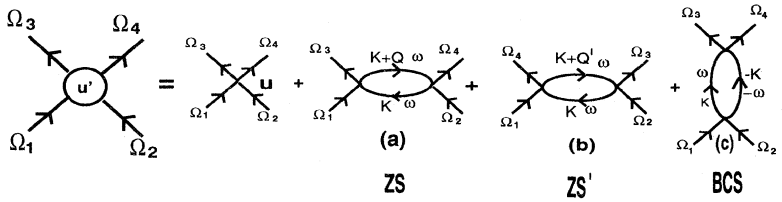

FIG. 9. The diagrams that renormalize the marginal quartic couplings for a $d=2$ circular Fermi surface. All external frequencies are chosen to vanish and all external momenta are on the Fermi surface at angles given by $\mathbf{\Omega}_i$. The couplings that are marginal at the fixed point (at tree level) have $\mathbf{\Omega}_3=\mathbf{\Omega}_1$ and $\mathbf{\Omega}_4=\mathbf{\Omega}_2$ (F coupling), or the above with $4\leftrightarrow 3$ (coupling $-F$, by Fermi statistics) or $\mathbf{\Omega}_2=-\mathbf{\Omega}_1$ and $\mathbf{\Omega}_4=-\mathbf{\Omega}_3$ (coupling V). All loop momenta must lie within a shell of width $d\Lambda$. In the first case (F), there is no momentum transfer at the left vertex ($Q=0$) of the ZS graph. No matter what direction we choose for K, the other loop momentum $K+Q=K$ will also lie in the shell. Thus there is no restriction on the angle of the loop momentum. However, the two propagators have the same ω, and the ω integral gives zero, since the two propagators have the same momentum and hence the same energy. In the other two graphs the loop angle is restricted to lie in a region of width $d\Lambda/\Lambda$. Figure 10 shows this for the ZS' graph, and Fig. 8 can be used to show this for the BCS graph if we replace the annuli of thickness Λ by shells of thickness $d\Lambda$. Thus none of the diagrams causes a flow of F at the fixed point. The flow of V receives no contribution from the ZS and ZS' diagrams because there is no correlation between incoming and outgoing momenta; both Q and Q' will be of order K_F, and the diagrams will be suppressed by $d\Lambda/\Lambda$. The BCS diagram, however, does cause a flow of V exactly as in $d=1$: the loop frequencies are equal and opposite, but the loop energies are equal due to time-reversal invariance.

external legs at zero frequency and on the Fermi surface ($k=\omega=0$), since the dependence on these variables is irrelevant. As for the angles, we chose $\mathbf{\Omega}_1=\mathbf{\Omega}_3$. Consider the ZS diagram in Fig. 9 given by the first integral in Eq. (380). Since $Q=0$ here, we know that, if K lies in the shell being eliminated, so does $K+Q$ for any direction of K. In other words, θ runs over the full range. However, this diagram vanishes for the same reason it did in the $d=1$ case: both poles in the ω plane are on the same side. Even if we put in a small momentum transfer $Q\ll\Lambda$ at the left vertex, it will not change anything. This is because the loop momenta are near $k=\pm\Lambda$, and it takes a minimum momentum transfer of order Λ to knock something from below K_F to something above it and vice versa. This is in accord with the general statement that the bare coupling has no singular limit at small ω or Q.[15]

Now consider the ZS' diagram. Due to the momentum transfer $\mathbf{Q}'$ of order K_F at the left vertex, not only is the magnitude of the loop momentum restricted to lie within the shell being eliminated, but also its angle is restricted to a range of order $d\Lambda/K_F$. This is clarified by Fig. 10. The dark circles now represent the thin slices being integrated. The intersection regions, of order $d\Lambda^2$, give us the allowed loop momenta for the β-function calculation. They lie in the shell and have the right momentum transfer K_4-K_1. Of the eight intersections, only the four marked ones satisfy the condition of being on opposite sides of K_F so that the ω integral survives. Since the ω integral gives a denominator of order Λ, the contribution to du is of order $(d\Lambda/\Lambda)(d\Lambda/K_F)$, so that the β function vanishes as we take the limit $dt=d|\Lambda|/\Lambda\to 0$.

It is also clear that, if $\mathbf{Q}'=0$ and $\mathbf{Q}\neq 0$ (which obtains if $\mathbf{\Omega}_4=\mathbf{\Omega}_1$), we can repeat the above argument with ZS$\leftrightarrow$ZS'.

You may check that in the field theory approach we would get a β function that went as Λ/K_F. Since we are ultimately going to send $\Lambda/K_F\to 0$, this will not matter in that limit.

Finally, for the same kinematical reason, the BCS diagram does not renormalize F at one loop. Consider Fig. 8 with K_3 and K_4 replaced by the two momenta in the

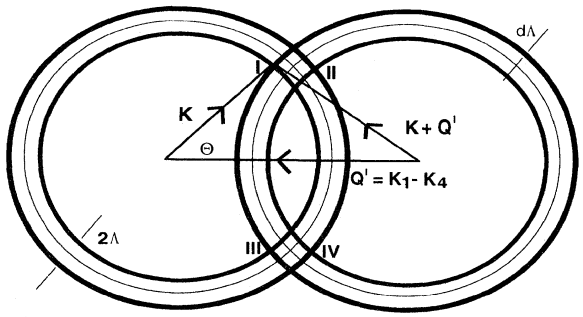

FIG. 10. Construction for determining the allowed values of loop momenta in ZS'. The requirement that the loop momenta come from the shell and differ by Q' forces them to lie in one of the eight intersection regions of width $(d\Lambda)^2$. Of these, four correspond to opposite signs of energies. But this is moot, since the extra power of $d\Lambda$ makes them irrelevant. (In the field theory approach, the suppression factor will be Λ/K_F instead of $d\Lambda/K_F$ and once again will be unimportant at the fixed point.)

[15]What if we choose to follow the field theory approach? The situation is exactly as in $d=1$. We must evaluate the one-loop graphs contributing to Γ, which is sensitive to whether we take the $Q\ll\omega$ or vice versa. The limit that gives a nonzero Γ ($\omega\ll Q$) makes a contribution of the form

$$\int d\omega\to\frac{f(\mathbf{K})-f(\mathbf{K}+\mathbf{Q})}{E(\mathbf{K})-E(\mathbf{K}+\mathbf{Q})}$$

where f is the Fermi function. As $Q\to 0$, this becomes a δ function at K_F and makes a contribution that is Λ independent and hence irrelevant to the β function.

BCS loop, K and $P-K$. In each annulus we keep just two shells of thickness $d\Lambda$ at the cutoff corresponding to the modes to be eliminated. The requirement that K and $P-K$ lie in these shells and also add up to P forces them into intersection regions of order $d\Lambda^2$. This means the diagram is just as ineffective as the ZS' diagram in causing a flow. Thus any F is a fixed point to this order.

C. The one-loop β function for V

Let us now look at the evolution of V. We choose the external momenta equal and opposite and on the Fermi surface. The ZS and ZS' diagrams do not contribute to any marginal flow for the same reason that BCS and ZS' did not contribute to the flow of F: since Q and Q' are of order K_F, they are kinematically suppressed by an extra factor Λ/K_F in the field theory approach and by $d\Lambda/K_F$ in the modern approach. But the BCS diagram produces a flow in either approach. We follow the modern approach. The flow is due to the following factors: (i) The loop angle can run freely over its full range because, no matter what value K takes in the shell being eliminated, $P-K=-K$ automatically lies in the shell. (ii) The ω and k integrals behave as in $d=1$ and produce a factor $dt=|d\Lambda|/\Lambda$.

We find

$$\frac{dV(\theta_1-\theta_3)}{dt}=-\frac{1}{8\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}V(\theta_1-\theta)V(\theta-\theta_3)\ . \tag{381}$$

This is an interesting example of the β function for the coupling *function*. Fortunately we can simplify the picture by going to angular momentum eigenfunctions,

$$V_l=\int_0^{2\pi}\frac{d\theta}{2\pi}e^{il\theta}V(\theta)\ , \tag{382}$$

and obtain an infinite number of flow equations, one for each angular momentum l:

$$\frac{dV_l}{dt}=-\frac{V_l^2}{4\pi}\ . \tag{383}$$

The flow tells us that the couplings V_l, marginal at tree level, are marginally relevant if negative, and marginally irrelevant if positive. The former is just the BCS instability. As for the repulsive case, if we integrate the flow we get

$$V_l(t)=\frac{V_l(0)}{1+t[V_l(0)/4\pi]}\ , \tag{384}$$

which is just the observation of Morel-Anderson (1962) that interelectron repulsion is logarithmically reduced if we develop an effective theory for the modes close to K_F.

If this analysis were to be repeated in $d=3$, the only difference would be that the BCS β function for $V(z)$ would be decoupled using the Legendre polynomials $P_l(z)$ with l odd. The decoupled equation would have the same form as Eq. (383) with the same implications.

D. Fixed-point structure at one loop

Let us take stock: the tree-level fixed point is characterized by two marginal functions F and V. The function F is marginal at one loop also, while V is marginally relevant in an infinite number of ways, one for each angular momentum l, if attractive, and marginally irrelevant if repulsive. It appears that, even if a single V_l is negative, the coupling constants run off to some other massive fixed point with a BCS gap. What if all the V_l's are positive? It appears that these couplings will renormalize to zero logarithmically and we shall end up with a fixed point characterized by F. This turns out to be incorrect, at least in principle, due to some irrelevant operators that cannot be ignored, as first pointed out by Kohn and Luttinger (1965). Here is the point. An irrelevant operator by definition is something that renormalizes to zero, not something you can set equal to zero at the outset without any consequences. Before it renormalizes to zero, it can modify the flow of the relevant couplings. Recall the case of u_0, the scalar coupling in $d=4$. Although it was irrelevant, adding it to the Gaussian fixed point generated a mass term r_0, which then grew rapidly. A very similar thing happens here: an irrelevant term produces a small negative BCS coupling, which then grows rapidly. This subtle issue is discussed in the next subsection.

We discuss here another irrelevant term, which does not destabilize the fixed point, but modifies our description of it.

Consider the sunrise diagram, Fig. 11. In mode elimination this diagram comes from taking two quartic terms and seeing how they feed back on the quadratic term. Though it is also of order u^2, it has two loops and may be ignored in the one-loop discussion we are having. But if one evaluates it, one finds it is irrelevant due to phase-space restrictions in the limit of small Λ/K_F.[16] Before we reach this limit, it can produce interesting effects.

Let us write its contribution as $-\Sigma(k\omega)\bar{\psi}\psi$, where Σ is called the *self-energy*. (It has no θ dependence due to rotational invariance.) If we Taylor-expand Σ as follows,

$$\Sigma(k\omega)=\Sigma(0)+i\omega\underbrace{(1-Z^{-1})}_{\frac{\partial\Sigma}{\partial i\omega}}+k\frac{\partial\Sigma}{\partial k}+\text{irrelevant pieces}\ , \tag{385}$$

we see that $\Sigma(00)$ is to be eliminated using a counterterm $\delta\mu$ of order u^2; $(1-Z^{-1})$ changes the coefficient of $i\omega$ to Z^{-1}; and $\partial\Sigma/\partial k$ changes the coefficient of k from K_F/m to $(1+(m/K_F)(\partial\Sigma/\partial k))K_F/m$. At the free-field fixed point we rescaled the field to keep the coefficient of both the quadratic terms fixed. Now we have seen that there is a manifold of fixed points parametrized by nonzero u.

[16]The details of the evaluation will not be provided here. The interested reader is asked to do the phase-space analysis.

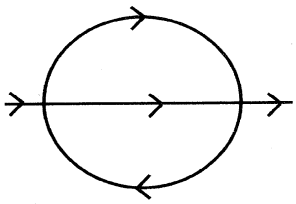

FIG. 11. The sunrise diagram contribution to self-energy. This has two loops and is suppressed by a factor Λ/K_F or, in the language of the $1/N$ expansion, by a factor of $1/N$.

In this case $i\omega$ and k may receive different contributions from eliminated modes and there is no rescaling that will keep both terms fixed, so we shall keep the coefficient of $i\omega$ fixed at unity, i.e., define

$$\psi' = s^{-3/2} Z^{-1/2} \psi \ . \tag{386}$$

This has two effects. The coefficient of k changes as follows:

$$\frac{K_F}{m} \to Z \left[1 + \frac{m}{K_F} \frac{\partial \Sigma}{\partial k} \right] \frac{K_F}{m} \equiv \frac{K_F}{m^*} \ , \tag{387}$$

which defines the effective mass m^*.

Next, the new quartic coupling is given by

$$u' = (u + \delta u) Z^2 \ , \tag{388}$$

where δu is the contribution we have already discussed. Question: does this modify the β function we calculated? Answer: not to order u^2, since Z deviates from unity at order u^2 and this produces changes of order u^3 in the equation above.

How about the fact that m^* is now moving as we renormalize? Upon looking at the kinematics of the sunrise diagram one can tell that, as Λ goes to zero, its contributions will vanish. Thus, although m will evolve to m^* in the early stages of renormalization, a fixed point characterized by some m^* will emerge asymptotically.

How can we have a fixed point in a theory where there is a nonzero dimensionful parameter K_F? The answer is that the fixed-point theory described above has no knowledge of K_F: the ZS graph gets all its contribution from the delta function at the Fermi surface, while the other two, which know about K_F, are suppressed by the factor Λ/K_F and vanish at the fixed point.

E. The Kohn-Luttinger effect

The flow of V_l was such that attraction led to instability, while repulsion meant downward renormalization to 0. There is no doubt the former is true, but the latter is incorrect in principle. The fault is not with the solution to Eq. (383), but with the equation itself. In deriving it, we ignored the contributions from the ZS and ZS′ channel graphs on the grounds that they were finite and down by powers of Λ/K_F, which made them unimportant in the limit $\Lambda/K_F \to 0$. But there is a surprise waiting for us if we go ahead and compute their contribution to the flow. As shown in Appendix B, the modified flow in $d=3$ is (upon setting many positive numbers independent of l and t to unity)

$$\frac{dV_l}{dt} = -\frac{V_l^2}{4\pi} - \frac{V^2(\pi)\lambda^{7/4}}{l^{15/2}[\lambda^{7/4} + l^{-7/2}]^2} \ , \tag{389}$$

where $V(\pi)$ is the backward-scattering amplitude in the BCS channel and $\lambda = \Lambda/K_F$. Notice that, as $\Lambda/K_F \to 0$, the second term vanishes as $(\Lambda/K_F)^{7/4}$, which is why we ignored it. (The strange power comes in because the intersection region scales in a special way when the momentum transfer is $\simeq 2K_F$, which is the region that dominates this piece; note the backward-scattering amplitude in the answer.) Why do we care about this piece?

Let us imagine that we are just beginning our renormalization. The input potential is the projection on the angular momentum l channel of some short-range potential. It follows that $V_l(t=0) \simeq e^{-l}$ as $l \to \infty$ in order that the sum of such coefficient times the $P_l(z)$ and all the derivatives of the sum with respect to z converge to a given analytic function $V(z)$. By contrast, the second term, at fixed λ, goes as $l^{-15/2}$ as $l \to \infty$. It is clear that, as soon as the flow begins, the exponentially small initial coupling $V_l(0)$ will very quickly be driven to negative values by the second term. Thereafter both terms will drive the instability.

This is the RG version of the famous Kohn-Luttinger argument, which is discussed at some length in Appendix B. The argument implies that, at $T=0$, the fixed point we studied always faces the BCS instability. However, one is still interested in this fixed point characterized by F. The reason is that T_c for the Kohn-Luttinger superconductor could be very low, or the l value for pairs absurdly high. Thus if we imagine a tiny nonzero temperature above this T_c, the instability will disappear. (For a recent survey of the Kohn-Luttinger effect see Baranov *et al.*, 1992). Recall from our analysis in Sec. III.A that a temperature $T = 1/\beta$ leads to an imaginary-time coordinate of range $0 \le \tau \le \beta$. In other words a quantum system infinite in space and at inverse temperature β is mapped by the path-integral method to a system in $d+1$ dimensions which is infinite in the spatial direction and of width β in the (imaginary) time direction. As we renormalize, the thickness in the time direction (in the sliding units) will get reduced by s just like the correlation length, and just unlike the momenta. Thus there is a flow to smaller β or larger T, the fixed point being $T = \infty$. But if the crossover is very slow, then in the interim the fixed point described by F will control the physics. Interestingly enough there are many real-world systems described in exactly that fashion. More on this in the section on Landau's Fermi-liquid theory.

VII. THE $1/N$ PICTURE, THE LEAP TO ALL LOOPS

So far we have followed the RG program to one loop in some detail. We must now see what happens at all loops. In general this would be a formidable problem. Luckily in the present problem a great simplification arises in the limit $\Lambda/K_F \to 0$. The presence of this small parameter allows us to relate the sum over all loops to the one-loop result.

Now a similar thing happens in theories where N, the number of species of fields, becomes very large. In this case, in the limit $1/N \to 0$, it is possible to sum over all loops, and the answer is expressed in terms of $1/N$. This is called the $1/N$ expansion.

We shall begin with a review of the $1/N$ approximation. It will then be shown that Λ/K_F plays the role of $1/N$. We begin as usual in $d=2$.

A. $1/N$ in $d=2$

Consider a ϕ^4 theory with action

$$S_0 = -\sum_{i=1}^{N} \phi_i^* \frac{k^2}{2} \phi_i - \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \phi_i^* \phi_i V_{ij} \phi_j^* \phi_j \ . \tag{390}$$

Many integrals are suppressed, and only the internal index is highlighted, since what we are about to say is independent of dimension. All we need to note is that there are N species of fields (or particles) and they have a quartic interaction V. The interaction has a factor $1/N$ in front of it to ensure that we have a nontrivial limit as $N \to \infty$. Note that, in the interaction vertex, if an i and a j come in, the same indices also exit, as shown in Fig. 12.

Let us look at a four-point function to one loop, as shown in Fig. 12. Among the one-loop graphs, only the first (ZS) is of the same order as the tree-level graph. This is because the extra factor of $1/N$ coming from the extra vertex is compensated by a sum over the loop index, which is free to take all values. This in turn was because the index i that came in went out at once, leaving the loop index l free to roam over all values. By contrast, the external indices have insinuated themselves into the loops in the other two diagrams (ZS′ and BCS), and there is no sum over the indices there. These graphs are then of order $1/N^2$ and hence suppressed by a factor $1/N$ relative to the tree graph. It is clear that the sum over iterated ZS loops ("the bubble sum") gives the leading behavior (in $1/N$) of the four-point function.

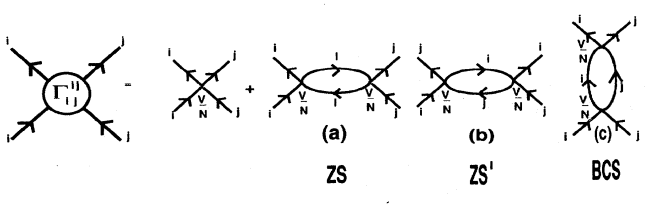

FIG. 12. Flow of indices in the $1/N$ analysis for the problem with an F-type interaction. The tree-level graph is of order $1/N$. The ZS graph is of the same order, since the sum over l pays for the extra vertex. The other two graphs are of order $1/N^2$. If we consider interactions of the V type, the BCS diagram and its iterates will be favored.

Note that the β function of this theory is completely given by the one-loop answer. This is because iterates of the ZS loop (the bubble sum) merely produce the nth power of $\ln\Lambda$ when iterated n times, while higher-order terms in β come from subdominant logarithms. This conclusion is also evident from Weinberg's (1993) discussion of the graphical content of the effective action.

Let us look at our theory now and consider a four-point function $\Gamma(\theta_2,\theta_1,\theta_2,\theta_1)$ in obvious notation with all external $\omega = k = 0$. We return to Fig. 12, this time using integrals over θ's in place of a sum over discrete indices. In the ZS graph, where the incoming "index" θ_1 immediately exits, there is no momentum transfer at the left (and hence right) vertex. Thus the loop angle runs over the full range 0–2π. In the ZS′ and BCS diagrams, on the other hand, there is a large momentum transfer Q' or large total momentum P (Figs. 8 or 10), which tell us that loop angles must lie within Λ/K_F of the external angles. In other words, the external angles have insinuated themselves into the loops and frozen the loop sum. Since the ratio of the ZS′ and BCS graphs to the ZS graphs is Λ/K_F, we expect that this ratio will play the role of $1/N$.

It should not be too surprising that we have a $1/N$ description available here: the noninteracting $d=2$ theory was written as an integral over the internal index θ, which labeled pseudo two-dimensional theories (with a phase space $dkd\omega$), one for each direction. But what should be the N assigned to this integral (rather than sum) over θ? Is it infinity, since there are infinite directions, or is it of order unity, since each has infinitesimal measure? Let us sharpen the analogy with $1/N$ analysis to answer this question. As a first step let us write the free-field action with all factors of K_F intact:

$$S_0 = \int_{-\Lambda}^{\Lambda} \frac{dk}{2\pi} \int_0^{2\pi} \frac{d\theta}{2\pi} K_F \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} [\bar{\psi}(i\omega - v^* k)\psi] \ . \tag{391}$$

Now we chop the angular integration into regions of width $\Delta\theta = 2\Lambda/K_F$, so that the annulus breaks up into $N = 2\pi K_F/2\Lambda$ cells labeled by an index i. The momentum of a point within a cell i is

$$\mathbf{K} = K_F \mathbf{\Omega}_i + k_i \mathbf{\Omega}_i + k_{\perp i} \mathbf{t}_i \tag{392}$$

$$\equiv K_F \mathbf{\Omega}_i + \mathbf{k}_i \ , \tag{393}$$

where $\mathbf{\Omega}_i$ is the unit radial vector at the center of cell i, $\mathbf{t}_i$ is a unit tangent, and k_i and $k_{\perp i}$ are radial and angular displacements from the cell center. We refer to the first piece of order K_F as the large momentum and the other as the small one. The measure per cell is

$$\int_{-\Lambda}^{\Lambda} \frac{dk}{2\pi} \int_{-\Lambda/K_F}^{\Lambda/K_F} K_F \frac{d\theta}{2\pi} = \int_{-\Lambda}^{\Lambda} \frac{dk}{2\pi} \int_{-\Lambda}^{\Lambda} \frac{dk_\perp}{2\pi} \equiv \int \frac{d^2k}{(2\pi)^2} \ . \tag{394}$$

In this notation

$$S_0=\sum_{i=1}^{N}\int \bar{\psi}_i(\mathbf{k}_i\omega_i)[i\omega-v^*k_i]\psi_i(\mathbf{k}_i\omega_i)\frac{d^2k_i d\omega_i}{(2\pi)^3} . \tag{395}$$

We now write down an interaction term,

$$S_F=-\frac{1}{K_F}\sum_{i,j=1}^{N}\int \bar{\psi}_j(\mathbf{k}_4\omega_4)\psi_j(\mathbf{k}_2\omega_2)F_{ij}\bar{\psi}_i(\mathbf{k}_3\omega_3)\psi_i(\mathbf{k}_1\omega_1) , \tag{396}$$

$$F_{ij}=F(\mathbf{\Omega}_i\cdot\mathbf{\Omega}_j) \tag{397}$$

$$\int=\int\prod_1^4\left[\frac{d^2k_i d\omega_i}{(2\pi)^3}\right]\delta(\omega_1+\omega_2-\omega_3-\omega_4) \times\delta^{(2)}(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3-\mathbf{k}_4)(2\pi)^3 . \tag{398}$$

Notice that the interaction conserves momentum: the fact that the index i appears once in a ψ and once in $\bar{\psi}$ means the large momentum is conserved. The small one is also conserved because of the explicit $\delta^{(2)}$ function. If we express all frequencies and momenta in units of Λ, it will be found that the only place where Λ appears will be in front of the interaction term in the form Λ/K_F, exactly playing the role of $1/N$.

Now, this interaction is not exactly the one we have dealt with so far; it forbids certain processes that were previously allowed. Consider a process in which the initial particles have momenta $\mathbf{K}_1\mathbf{K}_2$ from cells $i_1 i_2$, respectively, adding to a total $\mathbf{P}$. If we draw Fig. 8, we shall indeed find that $\mathbf{K}_3\mathbf{K}_4$ lie with the same intersection regions as $\mathbf{K}_1\mathbf{K}_2$, *but this need not mean they are in the same cells.* In other words, the intersection region can straddle more than one cell, if we imagine these cells permanently etched on the annulus. The old interaction would allow the full intersection region, while the new one would allow only the part of it in which the cell indices match in pairs. But notice that the difference between the two interactions shows up only in situations that are kinematically suppressed by a power of Λ/K_F. On the other hand, if we concentrate on just the sum over iterated ZS loops, with $\omega, Q \ll \Lambda$, it can be verified that the two give identical answers.

The careful reader will find one problem with the $1/N$ analogy. In the usual examples, u in Eq. (390) is held fixed as N varies. Here N is related to the cutoff Λ. As we lower the latter to increase the former, we must follow the flow of $F(\Lambda)=F(1/N)$. However, we have seen that there is no flow within the F couplings in the asymptotic region. Thus F_{ij} is essentially constant independent of N for large N.

In view of the above, we employ the following two-stage attack on the fermion problem:

Reduce the given problem in the full K space to a small-Λ (large-N) theory using the RG.

Solve the resultant theory using the smallness of $1/N$.

Consider now another type of coupling corresponding to V. In schematic form and the same notation used above,

$$S_V=\frac{\Lambda}{K_F}\sum_{ij}\int \bar{\psi}_{-j}(4)\bar{\psi}_j(3)V_{ij}\psi_{-i}(2)\psi_i(1) . \tag{399}$$

This interaction leads to a bubble sum (iterates of the one-loop graph) in the BCS channel. Here the coupling V grows as N increases, since there is a flow now. If we want to increase N keeping $V(N)$ fixed, we must start with weaker and weaker V's. If we do this, we can handle the BCS problem also by summing over bubble diagrams.

Let us hereafter assume that the BCS amplitudes V are absent in the spirit discussed earlier. Then we are left with iterated ZS loops and F's. The physics of this is Fermi-liquid theory, to be discussed in the next section.

B. $1/N$ in $d=3$

We have seen that the RG allows nonforward amplitudes to survive in $d=3$ as marginal interactions. If we divide the spherical Fermi surface into patches of size Λ^2 and label them by an index i, this will run over roughly $(K_F/\Lambda)^2$ values. The interaction F_{ij} will not be of the form of Eq. (398), since non-forward-scattering amplitudes are allowed. Let us, however, divide the possible interactions into a set involving just forward scattering and the rest. If we consider a forward-scattering four-point function Γ^{ij}_{ij} (with cell index conserved), it will be given as the bare vertex plus a sum of iterated ZS diagrams all involving only forward-scattering amplitudes. The insertion of a large-angle scattering anywhere will produce a factor Λ/K_F. [The suppression factor is *not* $(\Lambda/K_F)^2$ because kinematics only restricts the angle between 1 and 2 to be that between 3 and 4; the plane containing the latter is still free to rotate by the angle we called $\phi_{12;34}$.] If we consider instead Γ^{kl}_{ij}, which is nonforward, we find that it is given by just the tree-level coupling we put in the action. All loop corrections will be down by powers of Λ/K_F. If we consider any response function to a soft (low-wave-number) probe, these amplitudes will not enter the physics. If we compute the lifetimes of particles, these amplitudes will play a role, but phase space will again introduce powers of Λ/K_F. Thus the nonforward amplitudes live in a shadow world, perhaps large in magnitude but small in their effects.

As for the BCS amplitudes, we can find the stable state with a gap, using the $1/N$ expansion to limit ourselves to summing bubble graphs. Since nonforward amplitudes never enter the computation, this explains why it is permissible to use the reduced BCS Hamiltonian, in which only scattering within pair states is kept. (These amplitudes are just our V's, of course.)

Note that, as in any $1/N$ theory, the β function at one loop (which we have calculated) is all there is in the large-N or small-Λ limit.

C. Two-point functions at large N

Let us close by asking what happens to two-point functions at large N. Let us look at some of the graphs contributing to G^{-1}, shown in Fig. 13, if we use the one-loop fixed-point action Eq. (378). We see that all iterates of the tadpole are exactly canceled by the one-loop counterterm or fixed-point chemical potential $\delta\mu^*$: whenever we can draw one more loop, we can also use the counterterm which exactly kills it. The sunrise diagram, on the other hand, brings in new corrections. But this diagram is suppressed by $1/N$ and is ignored in the limit we are interested in. This means that, in this theory, *we know the exact self-energy of a particle for the given interaction F.* This will play a big role in Fermi-liquid theory, to which we shall turn our attention shortly. (One way to understand the above result is the following. When the shell thickness goes to zero, the limiting theory acquires particle-hole symmetry. We are quite accustomed to this symmetry's determining the chemical potential, even on a lattice. This will continue to be true even for nonspherical Fermi surfaces.)

The picture we developed in the large-N limit agrees with rigorous calculations of Feldman *et al.* (Feldman and Trubowitz, 1990, 1991; Feldman *et al.*, 1992, 1993), as explained to me by E. Trubowitz. According to these authors, the system always goes to a BCS state at $T=0$, and, if the BCS diagrams are eliminated, the rest define a convergent series with a finite radius of convergence.

Further details may be found in the references given.

VIII. LANDAU'S FERMI-LIQUID THEORY

Nearly four decades ago Landau (1956, 1957, 1959) attacked the problem of interacting fermions at very low temperatures $T \ll K_F$. Assuming the system evolved continuously from the noninteracting limit, he developed a phenomenological theory that proved very successful, for example, in the study of helium-3 (Pines and Nozières, 1963; Nozières, 1964; Leggett, 1975; Vollhardt and Wölfe, 1990; Baym and Pethick, 1991; Lifshitz and Pitayevskii, 1980.) The picture he arrived at, called Fermi-liquid theory, may be described in the terminology of this paper as the *fixed point described by F.* For many readers of Landau's work there was an element of mystery surrounding some of the manipulations. This had to be so, since he substituted for forty years of subsequent developments (the RG in particular) with his remarkable intuition.

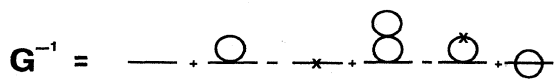

FIG. 13. Role of the counterterm (shown by a cross) in G^{-1}. The tadpole (order u) is canceled by the next diagram with just a cross. The value of the cross is found by first evaluating the tadpole and then arranging the cancellation. The next one with two loops is canceled exactly by the one following it, where the cross still stands for its order-u value. The sunrise diagram (last) will require a change in the counterterm, but this will be of relative order $1/N$ and hence ignorable in the large-N or small-Λ/K_F limit.

Following his work, a diagrammatic derivation of Fermi-liquid theory was provided by Abrikosov *et al.* and is described in their book (Abrikosov *et al.*, 1963). While the details were rather tedious, they did a lot to clarify where everything came from. I believe the approach developed here, using the RG, provides an even simpler route to Fermi-liquid theory, at least for those with a certain background.

So let us pretend we do not know what Fermi-liquid theory is and ask how we would arrive at it. (This paper specializes in $T=0$, and one expects the results to work also for the crossover region at low T with minor changes.) Using the RG developed here, we would find that, after repeated renormalization, we would have mapped the initial problem to one with $\Lambda/K_F \to 0$. Setting $V=0$ eliminates the BCS instability, i.e., implements Landau's requirement that there be no phase transitions, leaving us with a fixed-point theory parametrized by the marginal couplings m^*, $F(z\phi)$. The physics of this fixed point is Landau's theory. The excitations of this effective theory are the *quasiparticles* of Landau, in contrast to the "bare" particles created by the fields we began with prior to mode elimination. The fact that the quasiparticles have infinite lifetimes was established by Landau using phase arguments. In the present analysis, the lifetime terms, which appear as $O(\omega^2)$ terms in the self-energy, are irrelevant under the RG transformation. (At the free-field fixed point, they fall by s^{-1} under a factor s reduction of the cutoff.)

Is there any interesting physics in this limit $\Lambda/K_F \to 0$? If the cutoff is going to zero in laboratory units, are there any Feynman diagrams at all, or do we just read off all the physical scattering amplitudes from the vertices in the action, with no loop corrections, there being nothing left to run in the loops? In other words, is the full vertex function Γ the same as the (duly antisymmetrized) couplings F in the action?

The answer depends on whether we are looking at $\phi=0$ or $\phi\neq 0$, i.e., at processes in which the final pair lies or does not lie in the same plane as the initial pair. As for the latter, it is indeed true that tree-level amplitudes in the action would not be dressed by any loops due to the kinematics in $d>1$. For example, at one loop, the loop momentum would be restricted in magnitude and angle to size Λ and Λ/K_F, respectively. This, coupled with the fact that these diagrams have no singularities, would make them negligible, and we would have $\Gamma=F$.

For forward scattering, however, the iterated loops in the ZS channel, with only forward-scattering couplings (F) appearing in all the vertices, would have no restriction on the loop angle, which is why the graphs survived in the large-N picture. However, the *magnitude* of loop momentum k would still be bounded by Λ. Why would these graphs survive in the limit of vanishing cutoff? The

answer is that the integrand has a δ-function singularity at the Fermi surface (derivative of the Fermi function) if the external frequency transfer is zero, rendering the integral insensitive to Λ as long as it is nonzero.

Any attempt to introduce nonforward amplitudes $F(\phi\neq 0)$ into the iterated ZS loops would bring in a suppression factor, call it $1/N$ or Λ/K_F. Thus the forward and nonforward amplitudes do not mix. If we focus our attention on computing responses to soft probes $[(\omega Q)\ll\Lambda]$, we can ignore the nonforward F's. The resultant theory is just *Landau's Fermi-liquid theory.* The function $F(z,\phi=0)$ is called *the Landau parameter* $F(z)$.

Landau's $F(z)$ can be introduced in another equivalent way. Let us begin with our fixed-point theory (hereafter in $d=3$),

$$S=\int \bar{\psi}[i\omega-v^*k]\psi\frac{dk\,d\Omega\,d\omega}{(2\pi)^4}+\delta\mu^*\int\bar{\psi}\psi + \frac{1}{2!2!}\int u\,\bar{\psi}\bar{\psi}\psi\psi\ , \tag{400}$$

where u can contain nonforward amplitudes as well. We have seen that if

$$\delta\mu^*(\omega k\Omega)=\int_{-\Lambda}^{0}\frac{dk'}{2\pi}\int\frac{d\Omega'}{4\pi^2}F(\Omega'\cdot\Omega) \tag{401}$$

then there are no self-energy corrections in the limit we are in. Imagine now a state with a macroscopic number of (quasi) particles added to the ground state, so that all states up to momentum $k=r(\Omega)$ are occupied in the direction Ω. We can make such a state the ground state by modifying the Hamiltonian to

$$H'=H-\int\int\frac{dk\,d\Omega}{(2\pi)^3}\psi^\dagger v^*r(\Omega)\psi\ . \tag{402}$$

The action now becomes

$$S=\int\bar{\psi}[i\omega-v^*(k-r(\Omega))]\psi\frac{dk\,d\Omega\,d\omega}{(2\pi)^4} +\delta\mu^*\int\bar{\psi}\psi+\frac{1}{2!2!}\int u\,\bar{\psi}\bar{\psi}\psi\psi\ . \tag{403}$$

Let us now calculate the energy $\varepsilon'(k\Omega)$ (associated with H') of a particle in a state labeled by (Ωk). It is found from G^{-1}, which we can calculate exactly in the large-N limit, since we just need to evaluate the tadpole. This gives

$$\varepsilon'(k\Omega)=v^*(k-r(\Omega)) -\int\frac{d\omega'dk'd\Omega'}{(2\pi)^4}\frac{F(\Omega\cdot\Omega')}{i\omega'-v^*(k'-r(\Omega'))}-\delta\mu^*\ . \tag{404}$$

Doing the ω integral gives us $\theta(-k'+r(\Omega'))$, whereas the integral in $\delta\mu^*$ gives $\theta(-k)$. As a result, the energy ε of a particle, now measured with respect to H (which differs by v^*r from ε') is

$$\varepsilon(k\Omega)=v^*k+\int\frac{d\Omega'}{(2\pi)^3}F(\Omega\cdot\Omega')r(\Omega')\ . \tag{405}$$

It is evident that the integral involving F represents the interaction between the (quasi) particle in question and the rest. Consider now a state with $\delta n(k\Omega)$ quasiparticles at the point $(k\Omega)$, with $\delta n=-1$ if it is a hole. In terms of $r(\Omega)$,

$$\delta n(k\Omega)=\theta(k)\theta(r(\Omega)-k)-\theta(-k)\theta(k-r(\Omega))\ . \tag{406}$$

The energy of such a state, with reference to the ground state, follows from Eq. (405):

$$E[\delta n]=\sum_{k\Omega}v^*k\delta n(k\Omega) +\frac{1}{2V}\sum_{k\Omega}\sum_{k'\Omega'}\delta n(k\Omega)\bar{F}(\Omega\cdot\Omega')\delta n(k\Omega') +O(\delta n^3)\ , \tag{407}$$

where V is the volume, kept finite so we can do a sum rather than integral over momenta, and $\bar{F}$ is proportional to F, which is how Landau introduced his F.

Some texts (Mahan, 1981) devote some time to why Landau went on to keep the quadratic term in δn, and if he did, why he did not go to higher orders. The RG approach provides its own version of the answer. Both terms [coming from $\bar{\psi}(i\omega-v^*k)\psi$ and $\bar{\psi}\bar{\psi}\psi\psi$] are marginal, *whereas higher powers of ψ would bring in irrelevant operators.* In particular, the term with four powers of ψ competes with a term with just two, since the latter has an extra k or ω multiplying it, and these renormalize downwards under the RG transformation.

Of course we did not have to wait for this analysis to understand why Landau did what he did. If we go from Eq. (405), which gives the energy of an excitation, to the energy of all of the excitations, we find

$$\sum\to\frac{V}{(2\pi)^3}\int K_F^2dk\,d\Omega\ ,$$

$$\frac{E[r(\Omega)]}{V}=\frac{K_F^2}{(2\pi)^3}\left[v^*\int_0^{r(\Omega)}k\,dk\,d\Omega+\frac{1}{2}\int\int\frac{d\Omega}{(2\pi)^3}d\Omega'r(\Omega')F(\Omega\cdot\Omega')r(\Omega)\right]$$

$$=\frac{K_F^2}{(2\pi)^2}\left[v^*\int r^2(\Omega)\frac{d\Omega}{4\pi}+\frac{1}{2\pi^2}\int\int\frac{d\Omega}{4\pi}\frac{d\Omega'}{4\pi}r(\Omega')F(\Omega\cdot\Omega')r(\Omega)\right]\ . \tag{408}$$

We may view the above expression as representing the elastic energy of a membrane or rubber band representing the Fermi surface, with $r(\Omega)$ as the deformation parameter. (This equation may be found on page 54 of Nozières and Pines.) It is now clear that both terms are of the same order in the deformation. Haldane (in his lectures, yet to be published) emphasizes this aspect of Landau theory. He writes down the effective H as a *quadratic* function of some current densities obeying an algebra familiar in Conformal Field Theory. It is clear that we are discussing a solvable theory. The fact that only forward-scattering interactions $[F(z,\phi=0)]$ enter this theory is what makes it solvable and also endows it with additional symmetries. A very concrete application of Haldane's approach may be found in the work of Houghton and Marston (1992).

More recently Castro Neto and Fradkin (1993) used a coherent-state path integral to sum over the configurations of the Fermi surface.

This concludes the link between the present formalism and Fermi-liquid theory. Once we have the concept of the Fermi-liquid theory, there is no need for the RG. However, for the sake of those RG-minded readers who have followed all these arguments but are not familiar with Fermi-liquid theory, three sample problems will be discussed, not only to provide some instant gratification but also because each of them tells us something very instructive.

A. Landau theory for the masses

Our fixed-point theory parametrized by v^* and $u^* \simeq F$ has evolved from some bare theory with mass m, coupling U, and so on. In the RG approach, one does not usually attempt to reconstruct the bare parameters in terms of the final fixed-point quantities due to the unavoidable loss of information that accompanies mode elimination. There is, however, an ingenious argument due to Landau which does exactly that by relating m to m^* and F. The reason not all information about bare quantities is irreversibly lost is due, as always, to a symmetry, Galilean invariance being the operative one here. Let U be a unitary operator that acts on a state $|\psi\rangle$ and gives it an infinitesimal boost with velocity

$$\delta \mathbf{v} = \delta \mathbf{p}/m \ .$$

Under this active transformation, the energy of the eigenstate $|\psi\rangle$ changes as

$$\delta E = \langle U\psi|H|U\psi\rangle - \langle \psi|H|\psi\rangle \ . \tag{409}$$

Now Galilean invariance is the statement that *the boost affects only the kinetic energy of the particles and not the interaction energy.* Thus the response of H to this transformation is the same as in free-field theory and given by

$$U^\dagger H u = H + \mathbf{P}\cdot\frac{\delta \mathbf{p}}{m} + \cdots \tag{410}$$

where $\mathbf{P}$ is the total momentum operator. Thus to first order

$$\langle U\psi|H|U\psi\rangle - \langle \psi|H|\psi\rangle = \langle \psi|\mathbf{P}\cdot\frac{\delta \mathbf{p}}{m}|\psi\rangle \ . \tag{411}$$

Let $|\psi\rangle$ be a state containing one extra particle at the Fermi surface. Since the ground state has zero momentum, this is a state of momentum K_F in the chosen direction. Let the boost be in the same direction. The energy change according to the right-hand side of the above equation is

$$\delta E = K_F \frac{\delta p}{m} \ . \tag{412}$$

As for the left-hand side, the active transformation has three effects, which change the energy to *first order* in the boost:

(i) The quasiparticle momentum goes up by δp and, by the very definition of effective mass or Fermi velocity, its energy goes up by $K_F \delta p/m^*$.

(ii) The sea gets bodily shifted by $\delta \mathbf{p}$. This does not affect the kinetic energy of the sea, since the total momentum of the sea was zero.

(iii) The interaction of the quasiparticle with the shifted sea changes its energy as per Eq. (405), with $r(\theta)=\delta p \cos\theta$, where the angle θ is measured relative to the boost.

Adding all the pieces and equating the result to what was given in the previous equation, we get the famous relation

$$K_F \frac{\delta p}{m} = K_F \frac{\delta p}{m^*} + \delta p \int F(z) z \frac{dz}{4\pi^2} \ , \tag{413}$$

where $z = \cos\theta$. In terms of the dimensionless function

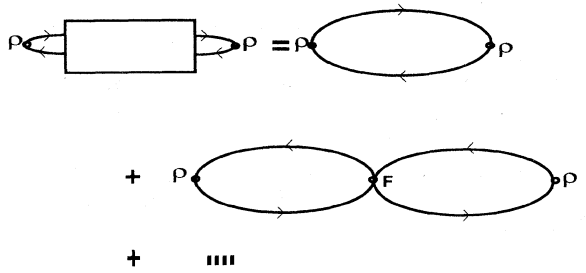

FIG. 14. The compressibility at small Q,ω, expressed as the correlation function of two densities. It is assumed we have renormalized down to a very small cutoff Λ, which is still larger than the probe momentum. The ρ's stand for the external probes coupling to the particle density. The first graph on the right-hand side is what one gets in free-field theory. The rest of the graphs are corrections due to interactions. If we try to introduce any large scattering amplitudes (these are marginal in $d=3$) into the graph (say, across one of the ZS bubbles), there will be a suppression by powers of Λ/K_F. Only the ZS bubbles survive the kinematical restrictions that all propagators lie within the narrow cutoff.

$$\Phi=\frac{m^*}{2\pi^2K_F}F=\frac{F}{2\pi^2v^*} \tag{414}$$

and the expansion in terms of Legendre polynomials

$$\Phi(z)=\sum_l \Phi_l P_l(z)\ , \tag{415}$$

we obtain

$$\frac{m^*}{m}=1+\frac{1}{3}\Phi_1\ . \tag{416}$$

B. Zero sound

Zero sound refers to natural oscillations in a Fermi liquid, resulting from interactions between the particles. We find it just as we would find the natural frequencies of an oscillator: by looking for poles in certain response functions. In our problem this means the density-density response function or compressibility.

Let us imagine an external probe $\phi(Q\omega)$ which couples to the density $\rho(Q\omega)$ of the fermions, producing density fluctuations. The compressibility χ is given by

$$(2\pi)^4\delta^{(4)}(0)\chi(Q\omega)=-\langle\rho(Q\omega),\rho(-Q-\omega)\rangle\ . \tag{417}$$

We are interested in the limit $Q,\omega\ll\Lambda$. We shall compute the correlation function by using the diagrammatic rules we have employed so far. The first few diagrams are given by Fig. 14. Because of the fact that the radial part of the δ function $[\delta(k-k')]$ that we pull out of our graphs differs from the traditional one $[\delta(k-k')/K_F^2]$ used in Eq. (417), we find

$$\frac{\chi(Q\omega)}{[-K_F^2]}=\int\frac{dk_1d\Omega_1d\omega_1}{(2\pi)^4}\left[\underbrace{\frac{1}{i\omega_1-v^*k_1}\frac{1}{i\omega_1+i\omega-v^*k_1-v^*Qz_{1Q}}}_{I(\omega_1k_1)}\right]$$
$$+\left[\int\frac{dk_1d\Omega_1d\omega_1}{(2\pi)^4}I(\omega_1k_1)\cdot\int\frac{dk_2d\Omega_2d\omega_2}{(2\pi)^4}I(\omega_2k_2)F(\Omega_1\cdot\Omega_2)\right]+\cdots \tag{418}$$
$$=\int\frac{Qz_{1Q}dz_{1Q}}{i\omega-v^*Qz_{1Q}}\frac{1}{4\pi^2}+\int\frac{Qz_{1Q}dz_{1Q}}{i\omega-v^*Qz_{1Q}}\frac{1}{4\pi^2}\int\frac{Qz_{2Q}dz_{2Q}}{i\omega-v^*Qz_{2Q}}\frac{1}{4\pi^2}F(\Omega_1\cdot\Omega_2)+\cdots\ , \tag{419}$$

where z_{iQ} is the cosine of the angle between Q and the direction of the loop momentum K_i.

Now, we would like to study the simplest problem of this kind and therefore would like to choose a constant F. Unfortunately, F, although not antisymmetric under the exchange of 1 and 2, still vanishes when the initial angles coincide. [See the nearest-neighbor example wherein $F\simeq(1-\cos\theta_{12})$.] We shall compromise and for once introduce spin. We shall assume that only up and down particles scatter, with a constant F_0. If we now look at Fig. 14, we see that the first loop gets an extra factor of 2 due to the spin sum. The second also gets only a 2, since spins at the vertex must be opposite. This restriction continues down the chain. Now we find that the series is geometric. The sum gives

$$\chi(Q\omega)=\frac{-2K_F^2I_0}{1-F_0I_0}\ , \tag{420}$$

$$I_0=\int\frac{Qz\,dz}{\omega-v^*Qz}\frac{1}{4\pi^2}\ , \tag{421}$$

where we have chosen to look at real (rather than Matsubara) frequency ω since we are looking for real propagating excitations.

Notice how we get an answer that is very sensitive to whether $Q/\omega\to0$ or vice versa. In the former case I_0 vanishes, whereas in the latter case it equals $-1/2\pi^2v^*$. This is the kind of sensitivity that plagues the four-point function Γ also, as alluded to earlier.

Clearly a pole occurs in χ when

$$\frac{1}{F_0}=\frac{1}{4\pi^2v^*}\int\frac{z\,dz}{s-z}$$
$$=\frac{1}{2\pi^2v^*}\left[\frac{s}{2}\ln\frac{s+1}{s-1}-1\right]\ , \tag{422}$$

where

$$s=\frac{\omega}{Qv^*}\ . \tag{423}$$

In terms of the dimensionless Φ_l introduced in Eq. (414),

$$\frac{1}{\Phi_0}=\left[\frac{s}{2}\ln\frac{s+1}{s-1}-1\right]\ . \tag{424}$$

For the solution to exist, we require $s>1$, i.e., the velocity of propagation ω/Q must exceed the Fermi velocity v^*.

We shall not discuss the extensive physics of this phenomenon. [For example, Mermin (1967) has showed that we shall always have a zero-sound mode, for any reasonable F.] The main point was to show that the narrow-cutoff theory has a lot of life in it.

C. Static compressibility

To find the equilibrium compressibility, we simply set $\omega \equiv 0$, $Q \to 0$ in the preceding calculation. This means that

$$\frac{Qz}{\omega - v^* Qz} \to -\frac{1}{v^*} . \tag{425}$$

The iterated integrals then simplify to the point that we no longer have to introduce spin or assume F is a constant. Going back to the spinless case, we find

$$\chi = \frac{m^* K_F / 2\pi^2}{1+\Phi_0} . \tag{426}$$

Had we computed χ in free-field theory we would have found

$$\chi_0 = \frac{mK_F}{2\pi^2} . \tag{427}$$

Thus

$$\frac{\chi}{\chi_0} = \frac{m^*/m}{1+\Phi_0} = \frac{1+\Phi_1/3}{1+\Phi_0} . \tag{428}$$

Now, no one will dispute that this is indeed the ratio of $\langle \bar{\psi}\psi\bar{\psi}\psi \rangle$ correlation functions in the fixed-point theory to those of the free-field theory. But in Landau theory one equates this to the ratio of compressibilities. This is not so obviously correct, and I thank N. Read for forcing me to clarify this point. The problem is this. Let us begin with the full path integral over the bare fields prior to the RG transformation with a coupling of the bare charge density to some external field A:

$$Z(A) = \int d\psi_0 d\bar{\psi}_0 e^{S(A)} , \tag{429}$$

$$S(A) = \int \bar{\psi}_0 (i\omega - E(K))\psi_0 + \frac{1}{2!2!} \int \bar{\psi}_0\bar{\psi}_0\psi_0\psi_0 U + A \int \bar{\psi}_0\psi_0 . \tag{430}$$

Now, it is certainly true that

$$\frac{\partial^2 \ln Z}{\partial A^2} = \langle \bar{\psi}_0\psi_0\bar{\psi}_0\psi_0 \rangle = \chi . \tag{431}$$

Suppose we now perform the RG transformation. In the process we rescale the field: $\psi_0 Z^{-1/2} = \psi$. (I apologize for using Z to denote two different things. Hereafter we shall only see the above definition, as the wave-function renormalization factor.) This means

$$A\bar{\psi}_0\psi_0 \to AZ\bar{\psi}\psi .$$

Now the partition function is preserved by mode elimination, and we can take its second logarithmic derivative with respect to A after the RG transformation to find χ. But this will give $Z^2\langle \bar{\psi}\psi\bar{\psi}\psi \rangle$, whereas we computed the operator without the Z's in what we called χ of the fixed-point theory.

So this is the mystery. The resolution lies in the fact that, besides the rescaling, a term of the form $A\bar{\psi}\psi$ is generated (in addition to what was already there) as we carry out the RG transformation, and this precisely cancels the effect just discussed. This is an example of a Ward identity based on charge conservation. [See Abrikosov *et al.* (1963) for a discussion.] Here is a *glimpse* of how it works with no numerical factors. Let us look at two graphs that cancel. (All loop momenta lie in the eliminated region and correspond to fast modes.) Take the sunrise diagram, Fig. 15(a), whose $i\omega$ derivative at $\omega = 0$ contributes to Z, the field rescaling factor:

$$i\omega \to i\omega \left[1 - \frac{\partial \Sigma}{i\partial\omega} \right] \equiv i\omega Z^{-1} . \tag{432}$$

Imagine routing the external momentum through the upper line. Taking the $i\omega$ derivative clearly squares that propagator. This is shown in Fig. 15(b), with the cross denoting the place where the second propagator joins the first.

Consider now the other phenomenon: generation of new terms. In the mode elimination scheme the coupling between A and $\bar{\psi}\psi$ can take place via the fast modes as shown in Fig. 15(c). Notice that this diagram coincides with that in Fig. 15(b) in the limit when the probe brings in zero momentum and frequency. Consequently the field rescaling effects (due to the self-energy diagram) precisely cancel induced terms effects (due to the vertex correction diagram). Although we took just a pair of dia-

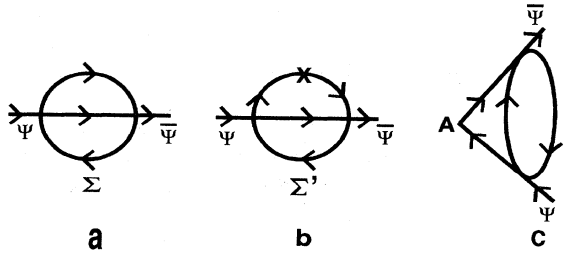

FIG. 15. Ward identity at work. The idea is to show that the coupling of the charge density to a field A is unaffected by interactions in the limit of vanishing probe frequency and momenta. (a) represents the renormalization of the quadratic coupling by the addition of $\Sigma(k,\omega)$. (The loop momenta are all fast.) If we expand Σ in a power series in $i\omega$, the first derivative will modify the quadratic piece by $\Sigma' = \partial\Sigma/\partial i\omega$. The field must then be rescaled to neutralize this. The graph in (b) represents Σ', the (external) frequency derivative of that in (a). The derivative introduces an extra propagator. The corresponding diagram represents the change in the field rescaling. Consider, however, the graph in (c), which reflects the fact that the slow fields can couple indirectly to the external field A via fast modes denoted by ψ_F. This equals the graph in (b) when external momentum and the frequency of the probe vanish. Consequently the field rescaling and modification of the coupling to the external field exactly cancel for these two graphs. The cancellation, however, is an exact statement. The physical idea is that, while interactions may cause one quasiparticle to break up into many, the total charge (to which the probe couples in the limit considered) is the same, since the dissociation conserves charge.

grams, the result is exact. It reflects the fact that, even though the quasiparticle can break up into many particles (so that its chance of being a single particle is reduced), the field can couple to the fragments now, and the total charge of the fragments (which is all the field couples to in the limit of zero frequency and wavelength) is that of the quasiparticle. (The careful reader will ask: what about the c-number term of the form A^2 in the action that comes from integrating fast modes? These contributions from the fast modes drop out as the external momentum and frequency vanish, which is the limit we are interested in.)

Thus a lot of Landau theory acquires its power due to the fact that not only are many quantities (like $\langle \bar{\psi}\psi\bar{\psi}\psi \rangle$) computable in the fixed-point theory, they directly correspond, with no intervening, unknown prefactors, to physical observables (like compressibility) due to Ward identities.

D. Notes for the experts

Here are some notes for readers who are familiar with the details of one or another of the ideas invoked earlier.

(i) In the diagrammatic treatment of Fermi-liquid theory, one organizes the graphs as follows (in the notation of Abrikosov *et al.*, 1963). First, one looks at the theory in the limit where the external transfers $Q,\omega \to 0$. The full four-point function is called Γ^k and corresponds to the limit $\omega/Q \to 0$. It is given as a sum of diagrams where the bare vertex is called Γ^ω, which corresponds to the limit $Q/\omega \to 0$, and is irreducible with respect to a pair of particle-hole lines which are singular at the Fermi surface. It is assumed that the bare vertex is analytic in its arguments, and the troublemakers, the particle-hole lines that produce all the singularities in the small (Q,ω) limit, are explicitly displayed.

In the RG approach, the bare vertex u contains all the *safe* modes, which include particle-hole lines, with at least one of them outside the cutoff. *The only lines shown explicitly in the cutoff theory are particle-hole lines, both within the cutoff, these being the modes yet to be integrated.*

(ii) Couplings corresponding to nonforward scattering, called $F(z,\phi \neq 0)$ in this paper, are very important for the study of lifetime effects and transport properties. A nice discussion of this may be found, for example, in Mahan's book (1981). Note in particular the discussion of the work of Dy and Pethick on page 947. These authors asked how the forward-scattering Landau parameters $F(z)$ may be extended to non-forward-scattering couplings and argued for a certain ϕ dependence based on symmetry under exchange. The additional factor of $\cos\phi$ they came up with is exactly what we find in Eq. (372) of this paper.

(iii) Even though F is a marginal coupling, there are no anomalous dimensions for the operators, in contrast to the fixed line in $d=1$ along which the fermion field and other operators have continuously varying dimensions. It is worth finding out if the interaction is a redundant operator in the sense of the RG.

(iv) Landau theory appears very much like a classical self-consistent theory. We understand this as coming from the large-N saddle point, which, like any saddle point, represents a form of classical limit.

(v) Some readers familiar with $1/N$ expansions may ask how the F's manage to change physical quantities like χ by factors of order unity, when one always thinks of interactions as producing changes of order $1/N$. [In other words, why are the dimensionless numbers Φ, appearing in, say Eq. (428), producing corrections of order unity?] The central feature of $1/N$ expansions is that loops of a certain kind are of order unity, since the loop sum pays for the extra factors of the interaction. If we compute a scattering amplitude, which exists only due to interactions, we get a term of order $1/N$, since that is the strength of any one coupling. This is a sum of a tree-level or bare term and iterated loops, all of the same order. In other words, the quantum corrections in the iterated loops are of the same order as the tree-level term. But if we compute something like the density-density correlation function (not often done in field theory), this is given by the polarization bubble, which is nonzero in free-field theory and hence of order unity in the $1/N$ series. Loop corrections to it (ZS graphs) are of the same order as well.

IX. NONCIRCULAR FERMI SURFACES: GENERIC

We now discuss a Fermi surface that has no special symmetries other than time-reversal invariance: $E(\mathbf{K})=E(-\mathbf{K})$. Once again we focus on $d=2$, discussing in passing the extension to $d=3$. A surface that meets these conditions is an ellipse and is depicted in Fig. 16.

The first step is to set up the RG transformation for the noninteracting problem. Since $|K|$ is no longer a measure of energy, we must draw contours of constant *energy* ε (measured from the Fermi surface) and retain a band of width Λ in either side of the Fermi surface. Thus our starting point is the action

$$S_0=\int_0^{2\pi}\int_{-\Lambda}^{\Lambda}\int_{-\infty}^{\infty}\bar{\psi}(i\omega-\varepsilon)\psi\frac{J(\theta\varepsilon)d\theta\, d\varepsilon\, d\omega}{(2\pi)^3}\ , \tag{433}$$

where θ parametrizes the Fermi surface, and J is the Jacobian for $(K_x K_y)\to(\theta\varepsilon)$. We shall expand

$$J(\varepsilon\theta)=J(\theta)+\varepsilon J_1(\theta)+\cdots \tag{434}$$

around the Fermi surface and keep just the first term; the rest will prove irrelevant.

The RG transformation is exactly as before, with ε in place of k. It is clear that higher-order terms in J renormalize to zero with respect to this RG transformation. The interaction term is

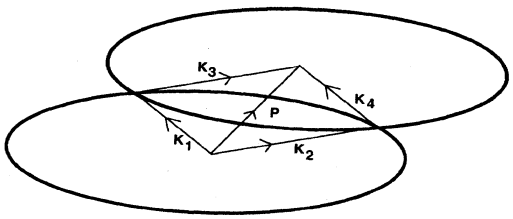

FIG. 16. A Fermi surface with time-reversal symetry but no rotational invariance. The dark line may be taken to represent a very thin shell left after a lot of renormalization. The figure shows that once again the set of initial vectors 1 and 2 coincides with the set of final vectors as the cutoff goes to zero. All the results of the circular case hold except for one thing: $F(\theta_1,\theta_2)$ and $V(\theta_3,\theta_1)$ are no longer functions of the differences of their arguments. The only instability is the BCS instability, which requires only time-reversal invariance. Indeed, the construction in the figure assumes this symmetry. The solutions it gives for $\mathbf{K}_3$ and $\mathbf{K}_4$ actually point from the surface to the center of the second ellipse instead of the other way around. However, time-reversal invariance assures us that, if we continue these vectors past the origin of the second ellipse, by an equal amount, we will hit the surface, so that $\mathbf{K}_3$ and $\mathbf{K}_4$ as shown are acceptable solutions. *In general, to find the final set using the initial set, we must draw the time-reversed version of the second surface and displace its center by* $\mathbf{P}$ *relative to the first.* This has not mattered so far due to the time-reversal invariance of the surfaces. If this condition were forgotten in the noninvariant case, one would erroneously conclude that when $\mathbf{P}=0$, the two surfaces coincide and the Cooper pairs can roam over all angles, leading to the BCS instability. The correct construction would show in this case that the two surfaces, even with $\mathbf{P}=0$, would intersect in only a few places of area $\simeq\Lambda^2$.

$$\delta S_4=\frac{1}{2!2!}\int\bar{\psi}(4)\bar{\psi}(3)\psi(2)\psi(1)u(4,3,2,1)\ , \tag{435}$$

where

$$\int\equiv\left[\prod_{i=1}^{3}\frac{J(\theta_i)d\theta_i d\varepsilon_i d\omega_i}{(2\pi)^3}\right]e^{-|\varepsilon_4|/\Lambda}\ . \tag{436}$$

A. Tree-level analysis

The analysis proceeds exactly as in the rotationally invariant case. As Λ (the cutoff in energy now) is reduced to zero (in fixed laboratory units) by mode elimination, we find once again that the set of initial momenta 1 and 2 must coincide with the final set 3 and 4. To see this, we must simply replace figures that have intersecting circles with, say, intersecting ellipses (Fig. 16) or whatever may be the shape of the Fermi surface. In a smooth cutoff only such couplings will not become exponentially small under renormalization.

Thus once again the marginal couplings at tree level obey θ_3 equal to θ_1 or θ_2 unless $\theta_1=-\theta_2$, in which case $\theta_3=-\theta_4$. (The figure shows the second possibility.) The tree-level amplitudes are labeled as before:

$$u(\theta_2\theta_1\theta_2\theta_1)=-u(\theta_1\theta_2\theta_2\theta_1)=F(\theta_2\theta_1)\ , \tag{437}$$

$$u(-\theta_3\theta_3-\theta_1\theta_1)=V(\theta_3\theta_1)\ , \tag{438}$$

but are *no longer* functions of the differences of their arguments.

B. Tadpole graph

We now see a new feature with the tadpole. Let us understand this without any reference to the RG. Suppose we begin with the action equation (435) and go to one loop. The tadpole graph, Fig. 5, makes the following contribution to the self-energy:

$$\Sigma(\theta\omega\varepsilon)=\Lambda\int F(\theta\theta')\frac{J(\theta')d\theta'}{(2\pi)^2}\equiv\varepsilon_0(\theta)\ . \tag{439}$$

Thus

$$G^{-1}=i\omega-\varepsilon-\varepsilon_0(\theta)\ , \tag{440}$$

and the Fermi surface has moved to $\varepsilon=-\varepsilon_0(\theta)$. *This generally involves a change in shape.*

Now even in the rotationally invariant problem the Fermi surface moves; turning on interactions changes the Fermi surface radius from $K_F^0=\sqrt{2m\mu}$ to $K_F(\mu)$ such that

$$\frac{K_F^2(\mu)}{2m}+\Sigma(u)=\mu\ . \tag{441}$$

In that case we can add a counterterm $\delta\mu^*=\varepsilon_0(\theta)=\varepsilon_0$ to the action to restore the old Fermi-surface radius. Recall that the action (schematic)

$$S=\int\bar{\psi}(i\omega-\varepsilon)\psi+\varepsilon_0\int\bar{\psi}\psi+\frac{1}{2!2!}\int u\bar{\psi}\bar{\psi}\psi\psi \tag{442}$$

was invariant under the RG at one loop and order u.

Now, it was pointed out that adding a counterterm to maintain the radius of the Fermi surface was not fine tuning, since it corresponds to maintaining a fixed density, which is experimentally viable. It was also pointed out that, if we did not add the counterterm, the system would not acquire a gap, as in ϕ^4 theory, but would simply move to the new radius defined in Eq. (441).

In the nonrotationally invariant problem, clinging to the old Fermi surface is certainly a case of fine tuning. There is no experimental way (such as sealing off the system) to preserve the detailed shape of the Fermi surface. We could add a *constant* $\delta\mu$ to keep the density (i.e., volume enclosed by the Fermi surface) constant, but we will not; we shall let the system find the true Fermi surface for the given μ and u. We can use the RG to determine the final Fermi surface as follows. We rewrite the initial action by adding and subtracting a presently unknown term $\varepsilon_0(\theta)$:

$$S=\int\bar{\psi}(i\omega-[\varepsilon+\varepsilon_0(\theta)])\psi+\int\varepsilon_0(\theta)\bar{\psi}\psi+\frac{1}{2!2!}\int u\bar{\psi}\bar{\psi}\psi\psi$$

$$\equiv\int\bar{\psi}(i\omega-\bar{\varepsilon})\psi+\int\varepsilon_0(\theta)\bar{\psi}\psi+\frac{1}{2!2!}\int u\bar{\psi}\bar{\psi}\psi\psi \tag{443}$$

and demand that this be a fixed point. To one loop and order u this means

$$\varepsilon_0^*(\theta)=s\left[\varepsilon_0^*(\theta)-\int_{-\Lambda/s}^{\Lambda}\int_0^{2\pi}\int_{-\infty}^{\infty}\frac{d\bar{\varepsilon}'d\theta'd\omega'J(\theta')F(\theta\theta')}{(2\pi)^3(i\omega-\bar{\varepsilon}')}\right], \tag{444}$$

which leads to

$$\varepsilon_0^*(\theta)=\Lambda\int\frac{d\theta'J(\theta')F(\theta'\theta)}{(2\pi)^2}. \tag{445}$$

To obtain this we have to do the ω integral in Eq. (444). Note that in that equation everything in the integral is evaluated at order u^0 due to an explicit F in it. This means that the propagator in the integral has an angle-independent $\bar{\varepsilon}'$.

Since we now have a fixed point, it must be true that we have found the correct Fermi surface. From the knowledge of ε_0^* we can reconstruct the new Fermi surface. In principle one could go order by order in this "renormalized perturbation theory." However, in the large-N limit, which appears here also, the one-loop answer gives the full self-energy correction or change in Fermi surface.

C. One loop at order u^2

If we do mode elimination as in the rotationally invariant case, we find once again that the ZS and ZS' graphs do not contribute to the flow of F for the same reason: either there is not enough momentum transfer to knock an internal line at $-\Lambda$ to Λ, or there is a kinematical suppression factor $d\Lambda/K_F$. The flow in the BCS channel is unaffected by nonrotational invariance. As long as $E(K)=E(-K)$ the BCS diagram, given by the third term in Eq. (315), will make a contribution and we shall obtain

$$\frac{dV(\theta_1;\theta_3)}{dt}=-\frac{1}{8\pi^2}\int_0^{2\pi}\frac{d\theta}{2\pi}V(\theta_1;\theta)V(\theta;\theta_3)J(\theta), \tag{446}$$

though we can no longer use rotational invariance to decouple this equation using angular momentum eigenfunctions. It is, however, possible to do a double Fourier expansion. This deserves further study analytically and numerically.

Weinberg (1993) has recently derived such a flow equation for superconductors whose Fermi surfaces obey just time-reversal invariance, using the notion of effective actions from quantum field theory. Besides the flow, his paper has a careful derivation and analysis of the effective action for superconductors. All the important properties of the superconductor may be derived from its effective action and from the notion of broken gauge invariance (Weinberg, 1986).

If we ignore the BCS interaction, we expect the Fermi liquid, $1/N$ etc., to work as before except for lack of rotational invariance. The F function will now be a function of two variables. We expect to find zero sound. We do not expect any simple relation between m and m^* due to a lack of Galilean invariance.

Finally, if we consider a Fermi surface without time-reversal symmetry, we can get rid of the BCS amplitudes even at $T=0$. (However, in drawing the analog of Fig. 8, we must draw the time-reversal inverted version of the second Fermi surface (displaced by $\mathbf{P}$), since the previous construction assumed that, if $\mathbf{K}$ is an allowed vector, so is $-\mathbf{K}$.

X. NONCIRCULAR FERMI SURFACES: NESTED

We are finally going to discuss spinless fermions on a square lattice at half filling. A specific model for this problem is the one with nearest-neighbor interaction,

$$H=H_0+H_I \tag{447}$$

$$=-\tfrac{1}{2}\sum_{\langle jj'\rangle}\psi^\dagger(j)\psi(j')+\text{H.c.}$$
$$+U_0\sum_{\langle jj'\rangle}[\psi^\dagger(j)\psi(j)-\tfrac{1}{2}][\psi^\dagger(j')\psi(j')-\tfrac{1}{2}], \tag{448}$$

where j labels sites on a square lattice and the subscript $\langle jj'\rangle$ on the sums means j' is restricted to be the nearest neighbor of j in the direction of increasing coordinates. [Thus if j is the origin (0,0), j' is restricted to be (1,0) or (0,1).] The chemical potential (found by opening up the brackets in the interaction term) is

$$\mu=2U_0, \tag{449}$$

where the factor 2 comes from the number of nearest neighbors.

At $U_0=0$, the half-filled system will once again be a perfect conductor, as can be seen by going to momentum states. Likewise at $U_0=\infty$ there will be a charge-density wave with more charge on one sublattice, say, the one whose x and y coordinates (which are integers in lattice units) add up to an even number.

Once again we shall focus not so much on the fate of this one model but rather on a class of models described by the same free-field fixed point and its perturbations. Let us therefore find the fixed point describing the noninteracting problem.

Let us take as the free-fermion dispersion relation

$$E=-\cos K_x-r\cos K_y, \tag{450}$$

which corresponds to the problem with unequal hopping in the two directions. The reason for the choice $r\neq 1$ will follow shortly. Notice that we expect to see a CDW state at large repulsion in the nearest-neighbor model even if $r\neq 1$ (since in this limit the hopping term is not the deciding factor; the interaction term is). Notice also that, for any r, E still has the symmetry

$$E(\mathbf{K}+\mathbf{Q}_N)=-E(\mathbf{K}),\quad \mathbf{Q}_N\equiv(\pi,\pi), \tag{451}$$

which is all we shall need. The chemical potential that gives rise to half filling at $U=0$ is $\mu=0$, which means that $E=\varepsilon$ and that the latter also changes sign when we add the nesting vector:

$$\varepsilon(\mathbf{K}+\mathbf{Q}_N)=-\varepsilon(\mathbf{K}) . \tag{452}$$

The Fermi surface for $r>1$ is sketched in Fig. 17. The dark line shows the Fermi surface, which now has two branches $\alpha=\pm 1$. Each point on the Fermi surface goes to another point on the Fermi surface upon adding $\mathbf{Q}_N$. This means that, if we shift the figure by $\mathbf{Q}_N$, the shifted figure (in the repeated zone scheme) will fit perfectly with the original like something out of Escher's drawings. If the momentum transfer is $\mathbf{Q}_N$, the analog of Fig. 10 will show the complete *overlap* of the two displaced surfaces rather than their intersection. Also shown in Fig. 17 are contours of energy $\varepsilon=\pm\Lambda$. This is where the modes to be eliminated under infinitesimal renormalization lie. Points at $\pm\Lambda$ are scattered to $\mp\Lambda$ upon transfer of $\mathbf{Q}_N$.

For nesting to take place, we need half filling and the symmetry in Eq. (452). The latter comes if we assume that hopping is always from one sublattice to the other in a bipartite lattice. Since hopping is usually just nearest neighbor to an excellent approximation, studying the effects of nesting once again does not constitute fine tuning.

Let us write down the action for the noninteracting problem. We shall use as the final coordinates ε and $\theta\equiv K_x$ together with a discrete index $\alpha=\pm 1$, which tells us which of the two branches we are on. Thus

$$S_0=\sum_\alpha \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{-\pi}^{\pi}\frac{d\theta}{2\pi}\int_{-\Lambda}^{\Lambda}\frac{d\varepsilon}{2\pi}J(\varepsilon\theta)\bar{\psi}_\alpha(i\omega-\varepsilon)\psi_\alpha , \tag{453}$$

where

$$J(\varepsilon\theta)=\frac{1}{\sqrt{r^2-(\varepsilon+\cos\theta)^2}} . \tag{454}$$

It is clear why we introduced $r\neq 1$: if it equaled unity, J would be plagued with (van Hove) singularities on the Fermi surface, and the expansion in ε would be impossible. Since nesting, and not van Hove singularities, is what we are interested in here, we shall study $r>1$. In this case the Fermi-surface value of the Jacobian is

$$J_0(\theta)=\frac{1}{\sqrt{r^2-\cos^2\theta}} . \tag{455}$$

Henceforth the subscript on J will be dropped.

Mode elimination of the action in Eq. (453) and the rescaling of fields and ε go as before to render the action

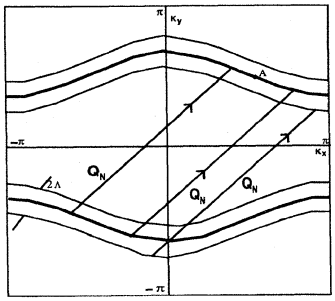

FIG. 17. A nested Fermi surface with $\mathbf{Q}_N=(\pi,\pi)$ and hopping anisotropy $r>1$. The filled states go from the origin to the dark lines, the two-branched Fermi surface. Any point on the Fermi surface (defined here by zero energy $\varepsilon=0$) goes to another on the surface upon addition of $\mathbf{Q}$, which reverses the energy. Consequently any point just below (above) the filled sea goes to a point just above (below) the sea, leading to the failure of perturbation theory at second order when a peturbation of momentum $\mathbf{Q}_N$ is introduced and to the flow of the coupling W through the ZS graph where the momentum transfer is $\mathbf{Q}_N$. The thin lines are equal-energy contours at $\varepsilon=\pm\Lambda$ and stand for the infinitesimal shells being integrated out in the RG program. The point A at $(\pi/2\pi/2)$ is privileged. It lies on the Fermi surface for all r and scatters into minus itself under the addition of $\mathbf{Q}_N$. The coupling for $A,-A\rightarrow A,-A$ is an F, a V, and a W and flows for both reasons. It is conjectured that this is probably why holes are found at these points (and two more, obtained by reflecting on the y axis) upon doping.

the fixed point.

As for the quartic term, there are now *three sets of couplings* that are marginal at tree level. Besides F and V, we also have

$$\begin{aligned} u[\theta_2+\pi,-\alpha_2;\theta_1+\pi,-\alpha_1;\theta_2\alpha_2;\theta_1\alpha_1] &= -u[\theta_1+\pi,-\alpha_1;\theta_2+\pi,-\alpha_2;\theta_2;\theta_1\alpha_1] \\ &\equiv W[\theta_2\alpha_2;\theta_1\alpha_1] , \end{aligned} \tag{456}$$

which corresponds to processes wherein the momentum transfer between 1 and 3 or 2 and 3 equals $\mathbf{Q}_N$. In this case, because of the nesting property of the Fermi surface, we are assured that the fourth momentum will lie on the Fermi surface if the first three do, and this, you recall, is the condition for the coupling to survive the tree-level RG transformation. Note also that W describes Umklapp scattering if particles 1 and 2 start out on the same branch and hop to the opposite one.

As for the one-loop β function, the F's do not flow, and V's have the usual flow given by

$$du[-K_3K_3-K_1K_1]=-\frac{1}{2}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{\text{shell}}\frac{d^2K}{(2\pi)^2}\frac{u[-K_3K_3-KK]u[-KK-K_1K_1]}{[i\omega-\varepsilon(K)][-i\omega-\varepsilon(-K)]} , \tag{457}$$

where "shell" refers to the shell being eliminated around both branches. The two contribute equally to give, in terms of V,

$$\frac{dV(\theta_3\alpha_3;\theta_1\alpha_1)}{dt}=-\frac{1}{2}\sum_{\alpha}\int\frac{J(\theta)d\theta}{(2\pi)^2}V[\theta_3\alpha_3;\theta\alpha]V[\theta\alpha;\theta_1\alpha_1]\ . \tag{458}$$

Since we do not have rotational invariance, we cannot separate this using the angular momentum variables. We can, however, use the double Fourier transform and reduce it to discrete coefficients, which will be coupled in their evolution.

Let us now look at the flow of $u[K_2'K_1'K_2K_1]$ where $K_i'=K_i+Q_N$:

$$du[K_2'K_1'K_2K_1]=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{\text{shell}}\frac{d^2K}{(2\pi)^2}\frac{u[K_2'KK_2K']u[K'K_1'KK_1]}{[i\omega-\varepsilon(K)][i\omega-\varepsilon(K')]} \tag{459}$$

where "shell" means that both $\varepsilon(K)$ and $\varepsilon(K')$ lie in the thin shells of width $d\Lambda$ near $\pm\Lambda$.[17] Due to the nesting property, two interesting things happen leading to a flow: If $\varepsilon(K)$ lies in the shell, so does $\varepsilon(K')=-\varepsilon(K)$. The ω integral never vanishes, since the poles always lie on opposite half planes. Doing the ω integral, we get, in terms of W,

$$\frac{dW[\alpha_2\theta_2\alpha_1\theta_1]}{dt}=-\int\sum_{\alpha}W[\alpha_2\theta_2\alpha'\theta']W[\theta\alpha\theta_1\alpha_1]\frac{J(\theta)d\theta}{(2\pi)^2}\ , \tag{460}$$

where $(\alpha'\theta')$ refers to K'.

To get a feel for this problem, let us evaluate the nearest-neighbor interaction on the present Fermi surface to obtain

$$W[\theta_2\alpha_2\theta_1\alpha_1]=-U_0\left[\sin^2\left[\frac{\theta_1-\theta_2}{2}\right]+\frac{1}{2}\left[1-\frac{\cos\theta_1\cos\theta_2}{r^2}-\frac{\alpha_1\alpha_2}{r^2}\sqrt{r^2-\cos^2\theta_1}\sqrt{r^2-\cos^2\theta_2}\right]\right] \tag{461}$$

and

$$V[\theta_3\alpha_3\theta_1\alpha_1]=U_0\left[\sin\theta_1\sin\theta_3+\frac{\alpha_1\alpha_3}{r^2}\sqrt{r^2-\cos^2\theta_1}\sqrt{r^2-\cos^2\theta_3}\right]\ . \tag{462}$$

It is readily verified that W always has the same sign for all values of its arguments—opposite to that of U_0. Specializing to the repulsive case, it is clear from the above equation that $d|W|/dt>0$. Thus we have proven the instability for this initial condition. (This is a weak-coupling argument. To describe the nearest-neighbor problem in the small-Λ theory it is not enough simply to restrict the full interaction to within the cutoff; we must take into account induced terms and renormalization due to elimination of modes. These, however, are higher-order effects.) As for some other interaction, if it has any overlap with this direction we have found, it will be unstable. Note that, in contrast to the rotationally invariant problem, we have not explicitly displayed an infinite number of unstable directions (one for each l).

Unlike the rotationally invariant case, the present problem has many open questions. Here is a list of some of the more important ones.

(i) The study of flow equations [Eqs. (458)–(460)] is a very important follow-through to the work presented here. Calculations are being performed with Murthy in which the couplings from the nearest-neighbor model [Eqs. (461) and (462)] are taken as initial conditions. In studying the flows, it is important to remember that there are some couplings that flow for more than one reason. An example is when the initial momenta are $\mathbf{K}$ and $-\mathbf{K}$ and the final ones are $\mathbf{K}'$ and $-\mathbf{K}'$, where $\mathbf{K}'=\mathbf{K}+\mathbf{Q}_N$. This coupling is equal to a V and a W. The general idea is to run the flow until the cutoff is small and then solve the theory by summing over diagrams that survive in the limit of vanishing Λ (our $1/N$). Such a study shows the CDW state for repulsive nearest-neighbor coupling and more exotic order if more interactions are added.

(ii) Although the $1/N$ formalism was discussed in connection with Fermi liquids with a rotationally invariant Fermi surface, it applies to all problems discussed here. One can always reduce the cutoff (keeping track of the evolving couplings) and then use the smallness of Λ to do a sum over diagrams that dominate the $1/N$ expansion. For rotationally noninvariant problems, however, this can be complicated by the constant motion of the Fermi surface as we renormalize, even if we keep its volume constant by modifying the chemical potential. In other words, as the modes are eliminated, the new Fermi surface and new one-particle energies must be used in defining contours of constant energy and choosing modes for the next round of elimination. At the one-loop level considered here, this was a nonissue.

(iii) It is important to consider the problem just below half-filling. Here we expect that there will be an initial

[17] We have defined W with $K_3=K_1'$, and for this choice only the ZS diagram contributes to flow. Had we reversed the role of K_3 and K_4 in the definition of W, an extra minus sign would have entered its definition, and the ZS' (rather than ZS) diagram would have contributed to its flow. The latter would also have had a minus sign relative to the ZS diagram.

growth of interactions as we lower the cutoff, which will then freeze once the cutoff is comparable to the deviation of the Fermi surface from nesting. In this case one must see whether, in the meantime, an attractive interaction is generated in the BCS channel. If so, this coupling will continue to grow, since it does not rely on nesting to do so. This will be yet another test of the notion that attraction can lie hidden in models that started out repulsive.

(iv) Another topic worthy of further study is the coupling $u(-A, A, -A, A)$, where $A=\frac{1}{2}\mathbf{Q}_N=(\pi/2,\pi/2)$ and lies on the Fermi surface for any r; see Fig. 17. [The present remarks apply equally well to the point $A'=(\pi/2,-\pi/2)$.] This coupling is a V, a W, and an F. It receives flow contributions from the ZS′ and BCS diagrams and is the forward-scattering amplitude which will control particle self-energies. Is the distinguished nature of this point (from the point of view of the RG) related to why so many investigators (Trugman, 1988; Sachdev, 1989; Elser *et al.*, 1990; Boninsegni and Manousakis, 1990) find holes occurring at the point A when the half-filled systems are doped? One could pick the initial value for the flow of the coupling generated by the nearest-neighbor interaction, follow the flow, and look at the effective theory at very low energies to see if an answer comes out.

(v) Consider the problem on a square lattice without nesting, say, because of hopping within the same sublattice. At very large nearest-neighbor repulsion, we can see that a CDW will result, with more particles in one or the other sublattice. However, there will be no instability at infinitesimal repulsion. One expects from continuity that the transition will take place at small coupling for small nesting violations. One should then see this phase transition at weak coupling from the RG.

(vi) Notice that our analysis depended on the nested Fermi surface. While the Fermi surface was nested in the absence of any interaction, do not the interactions cause it to move? Will not the shape change from perfect nesting, even if we change the chemical potential to sit at half filling? What happens to the CDW instability then? First of all, this question does not affect our one-loop calculation, which uses the zeroth-order propagators with their zeroth-order formula for ε. Whether or not the nested surface will stay nested at higher orders in the interaction is an open question, which could control the higher terms and thus decide the ultimate destination of the flow. It will not, however, change the fact that the free-field fixed point is unstable, since that has been established close to the fixed point by our one-loop calculation. However, if the CDW instability is really to take place, the flow must keep going until we reach a fixed point with a gap. Now, it is clear in coordinate space that, at strong coupling in a bipartite lattice, there will be a CDW. For this to come out of the RG, nesting must be preserved as we renormalize. This, however, is a conjecture and has not been proven.

Another interesting question is the following: if we set the hopping coefficient $r=0$, we seem to decouple the chains. Will we then get a Luttinger liquid? No. This is because the interaction terms couple the various decoupled chains. The exact cancellation that took place between the BCS and CDW instabilities in a one-chain model with just one coupling will not repeat itself anymore.

XI. NON-FERMI LIQUIDS IN $d=2$

So far we have seen two means by which the Fermi liquid could be destroyed: BCS and CDW instabilities. In both cases, the flow came about because *individual* Feynman diagrams had logarithmic singularities. Thus the perturbation series had zero radius of convergence. (There were essential singularities of the form $e^{-1/u}$ in, say, the CDW order parameter.) If perturbation theory can tell us about the instabilities, why follow the RG route? The answer, from the $d=1$ example, is that even if individual diagrams diverge, it is possible for the β function to vanish, producing novel scale-invariant behavior. We are looking for such a state in $d=2$. I see no evidence for it if the Fermi surface is spherical, the coupling is weak, the input interaction is short ranged, we work in an infinite volume from the start. In particular, I have examined, together with Ruckenstein and Schulz, the channel analyzed by Anderson (1990), in which the incoming particles were of opposite spin at the same momentum on or near the Fermi surface. We found that there was no flow in this coupling as the cutoff went to zero. Indeed there was no singularity in the diagram when the external (Euclidean) frequencies vanished. Setting them to nonzero values did not help. Of course, setting them equal to some *real* frequency did cause singularities, but these correspond to propagating modes and not instabilities of the ground state. (Recall that the usual instabilities were seen at zero external frequency.) But it must be pointed out that this is not at variance with Anderson's arguments, which rely very much on doing things in a finite volume and then carefully taking the infinite-volume limit. Unfortunately the RG in a finite volume is not an easy prospect, and we were hoping (despite Anderson's cautionary note about going to infinite volume too quickly) that if the effect showed up in our calculation, it would be an additional corroboration to Anderson's argument, but within the standard infinite-volume machinery. Engelbrecht and Randeria (1991), who studied this problem in the low-density expansion, did not find any instability.

We must therefore relax some of the above conditions. As mentioned above, Anderson dropped the infinite-volume condition. Another possibility (Varma *et al.*, 1989) is that at strong coupling a new possibility, the *marginal Fermi liquid,* which has impressive phenomenological success, arises. The present weak-coupling analysis has nothing to say about it and surely cannot exclude it. For example, it is possible for the bare coupling that enters the action of the narrow cutoff theory [or Γ^{ω}

in the Abrikosov *et al.* (1963) treatment] to be singular. Earlier it was stated that this would not happen because the bare couplings were obtained from the input parameters by integrating our safe modes. But this only assures us that the individual diagrams that add up to give the bare coupling are finite. It is certainly possible for the infinite sum to diverge beyond some maximum coupling. This does not contradict Landau's analysis, since the basic assumption that the physics in question is a continuation of the noninteracting problem is invalid.

Another possibility is that, even before mode integration, the input coupling is singular. Stamp (1992, 1993) has taken a pragmatic approach and considered the effect of singular interactions, setting aside the question of their origin. Let us consider the Coulomb interaction in this light. We cannot simply say that it gets screened; this is a picture that makes sense when a subset of diagrams in standard perturbation theory are resummed in a certain way to produce the screened propagator for the Coulomb potential. The bookkeeping is different in the RG: the coupling that goes into the action has not been screened by particle-hole pairs at the Fermi surface. It is easy to see (within the sharp-cutoff scheme) that at any stage $Q \ll \Lambda$ is unscreened. The correct procedure is to follow the evolution of the bare coupling as the modes are eliminated and see where it ends up when no more integration is left over. It is shown in Appendix A that the final potential is screened. (This is a smooth-cutoff version of a sharp-cutoff calculation devised with Murthy.) This analysis, however, assumes that the fermion propagator has the standard Fermi-liquid theory form. Ideally we should let the fermion propagator also evolve as modes are eliminated and see if we still end up with a screened interaction. This was recently done (Houghton *et al.*, 1993).

Now there *are* concrete examples of non-Fermi-liquid behavior if we are willing to consider impurity problems. Consider, for example, the example provided by Affleck and Ludwig (Ludwig and Affleck, 1991a, 1991b, 1992) from the Kondo problem. More recently Perakis *et al.* (1993) have given another example from the Kondo problem which shows non-Fermi-liquid behavior for a range of parameters.

Although the search for non-Fermi liquid did not yield anything at weak coupling, it is a worthwhile goal, since the copper-oxides seem to call for something different. As pointed out by Anderson, these may not be connected to the Fermi-liquid fixed point. Rather than reach the novel fixed point from the Fermi liquid, one could attempt writing down different fixed points. They may require additional fields besides fermions, e.g., gauge boson (Ioffe and Kotliar, 1990; Lee and Nagaosa, 1990; Polchinski, 1993). In a strong-coupling field theory the low-energy physics may bear no simple relation to the microscopic theory. Another route is to start with one or more one-dimensional systems, which can have Luttinger liquid behavior, and couple them perturbatively (Wen, 1990; Schulz, 1991; Di Castro and Metzner, 1992). Finally, one can study the problem in $1 \leq d \leq 2$ (Bares and Wen, 1993; Castellani *et al.*, 1993).

XII. SUMMARY AND OUTLOOK

The main aim of this paper was to find a way to apply the RG methods to interacting nonrelativistic fermions, in particular to understand the various instabilities for weak perturbations. In particular we wanted to see if the system remained gapless. Since the RG was so successful in dealing with critical phenomena, it was decided to follow a path relying heavily on analogy to this prior application. Since gapless systems correspond to critical systems, the idea was to use the language of fixed points and their perturbations.

We started with a brief historical review, from the original formulation of the RG for use in field theory to the modern approach pioneered mainly by Kadanoff and Wilson. It was pointed out that, while the RG has always expressed the invariance of the theory under a change in cutoff followed by a suitable change in the parameters, the emphasis has shifted from viewing the cutoff as an artifact to be sent to infinity (where it belongs in a continuum theory like quantum electrodynamics) to viewing it as a dividing line between interesting and uninteresting degrees of freedom (the slow and fast modes), even in a problem where the cutoff is finite to begin with.

In Sec. II we discussed the charged (complex) scalar field in four dimensions. It was shown how, if one wanted to study physics at long distances, one could work with just the "slow" modes in the functional integral and how the unwanted fast modes were to be eliminated and the new couplings deduced. The Gaussian fixed point was studied in detail and the flows to one loop were deduced. It was shown how, when the cutoff is reduced to very small values, i.e., when the phase space is reduced to a tiny ball, the coupling functions reduce to a handful of coupling constants, which were just the first few terms in the Taylor series of the coupling functions about the origin. A comparison was made between the modern way to find the flow equations (β functions), using mode elimination, and field theory methods (trying to get rid of cutoff dependence by proper choice of bare parameters). It was shown that, although the two had very different bookkeeping schemes, they gave the same answer for relevant and marginal couplings.

In Sec. III the functional integral method for fermions was introduced. First a few simple problems in the thermodynamics and dynamics of fermionic oscillators were solved using operator methods. Then the Grassmann integral formulation was introduced and the same results were regained. The rules for calculating correlation functions (Wick's theorem for fermions) were derived and used in the calculations.

The stage was now set for dealing with the nonrelativistic fermions. The strategy would be the following. We would first start with noninteracting fermions and

write a Hamiltonian that faithfully described the physics near the Fermi surface, i.e., within a cutoff Λ on either side of it. The logic was that the questions we were interested in were decided by these modes, at least at weak coupling. The corresponding functional integral would then be written. A mode elimination process that reduced the cutoff and rescaled momenta and fields and left the action invariant would be found. Given the RG and its fixed point, it would then be possible to classify the perturbations as relevant, irrelevant, or marginal. While this closely paralleled the scalar field problem, one major difference was anticipated: since we are renormalizing towards the Fermi surface and not the origin, the remaining phase space will be infinitesimal perpendicular to the Fermi surface, but of fixed size in the tangential direction. Thus the fixed point would be characterized by coupling functions that depended nontrivially on the angles used to parametrize the Fermi surface. The case $d=1$ was clearly exceptional, since the Fermi surface consisted of just two points. This made the problem very similar to continuum field theories wherein the phase space for bosons and fermions is a ball centered at the origin.

The study was limited to spinless fermions. At this point the reader can surely see that the inclusion of spin really is straightforward. Since all the interesting flows were due to special properties of the shape of the Fermi surface, and since spinless fermions could display these shapes, they were the clear pedagogical choice. To keep the discussion concrete, a nearest-neighbor spinless fermion problem at half filling was frequently invoked. The question was whether it developed a CDW gap at arbitrarily small repulsion.

The RG scheme worked remarkably well in the $d=1$ warmup, described in Sec. IV. We took two slices (of width 2Λ) near each Fermi point (L/R) and found a mode elimination scheme that left the action of the free theory invariant. We then turned on a quartic interaction. At tree level it was found that only the frequency and momentum ($k=K-K_F$)-independent part of this coupling was marginal. The rest were irrelevant, as were all couplings with six or more fields. The marginal coupling did depend on the discrete internal index L/R but, due to the Pauli principle, reduced to just one independent number. We had to go to one loop to resolve the fate of this marginal coupling. The β-function route avoided the CDW and BCS mean-field instabilities by playing them against each other and giving the correct answer: a scale-invariant system, the Luttinger liquid. (This correctly describes the nearest-neighbor model, which can be solved exactly.) We also discussed the quadratic terms that were induced by the quartic term upon mode elimination. These meant the Fermi surface was moving to take into account the interactions. There were two options. We could add a counterterm, determined order by order, to keep K_F fixed, or we could let the surface move to the new K_F. The latter would be found by doing renormalized perturbation theory: we split the original μ into two parts, one which gives the correct K_F in the propagators, and a counterterm that keeps it there. The division would be revealed to us order by order. It is a remarkable fact that Fermi systems, unlike Bose systems, can remain critical for a range of μ; they do this by using the fact that their fixed point is characterized by a whole surface, which can wiggle around as interactions are turned on. This idea applies with minor modifications in higher d, and we shall not speak of it further.

The scheme was then extended in Secs. V and VI to rotationally invariant Fermi surfaces in $d=2$ and $d=3$. In $d=2$ an annulus of width 2Λ and radius K_F was used. The RG transformation was essentially the same as in $d=1$ except for the fact that we had an integral over the "internal" variable θ rather than a sum over the Fermi points L and R. At tree level only two couplings, $F(\theta)$ and $V(\theta)$, which corresponded to forward scattering and Cooper-pair interactions, survived. They had no dependence on external k or ω. Going to one loop, we saw that there was no flow of F. This was because of the kinematics of $d>1$. There was, however, flow in V. We decoupled the flow into an infinite number of equations, one for each angular momentum l. The flow was marginally irrelevant for repulsion [corresponding to the findings of Morel and Anderson (1962)] and marginally relevant for attraction, corresponding to the BCS instability. We then formally set $V=0$ and identified F with Landau's F function. Thus the RG led us automatically to Fermi-liquid theory. The situation in $d=3$ was essentially the same, except for one significant difference: while in $d=2$ only forward-scattering amplitudes survived, in $d=3$ nonforward ones did also. However, the latter did not have any effect on low-energy, low-momentum transfer physics.

Having identified the Fermi liquid as the fixed point of our RG, we next asked why it is solvable, i.e., why, despite the quartic interactions, one is able to calculate many response functions. It was pointed out that, as we eliminate modes, a $1/N$ expansion emerges, with $N=K_F/\Lambda$. Landau theory is the $N=\infty$ limit. (By using certain collective coordinates, this limit can be made to correspond to a saddle point which gives the exact answer.) This is the first example in which N really is large in the original problem. However, a two-stage assault was needed to ensure this: first use the RG to bring Λ/K_F to a small value and then use $1/N$. Of course the $1/N$ is not a new way to do Landau theory, but just a new way to understand it. It has always been known that bubble graphs (ZS loops) dominate the low-momentum response. However, after the application of the RG there was nothing left but ZS bubble graphs due to the kinematics. In the bubble graphs the loop angle ran over all values. However, these need not have amounted to much, since the width of k integration was going to zero. But the graphs did survive, since the integrand had a δ-function singularity on the Fermi surface, and the integral was oblivious to the cutoff. (This is why the graph

survived but did not contribute to the β function, which probes the sensitivity to cutoff.)

An RG version of the Kohn-Luttinger instability was given. The details were relegated to Appendix B.

We briefly looked at three effects in Fermi-liquid theory; each taught us something and also gave the readers familiar with RG but not Fermi-liquid theory some instant gratification for their efforts.

By embedding Fermi-liquid theory in the framework of the RG and $1/N$ we not only automate the process that Landau had to finesse with his genius, we also prepare ourselves better to study variants of the problem he attacked, such as the problem with impurities, which is a big field (Lee and Ramakrishnan, 1985).

We then studied generic nonrotationally invariant problems. We found that we were led to F and V, which no longer depended on just the difference between their arguments. The flow equations for V were derived. By going to a Fermi surface with no time-reversal invariance, we could eliminate the BCS instability and have a real Fermi-liquid theory at $T=0$. The new feature here was the changing shape of the Fermi surface as we renormalized.

Moving on to the case of a nested surface in $d=2$, we found that a new coupling W, corresponding to momentum transfer $\mathbf{Q}_N=(\pi\pi)$, survived at tree level. At one loop this coupling began to flow. For the case of the perturbation corresponding to the nearest-neighbor interaction (truncated to modes within the cutoff), we saw the CDW instability. It was pointed out that the points $(\pm\pi/2\pm\pi/2)$ where holes seemed to appear upon doping the half-filled system are exceptional from the point of view of the RG: they are forward-scattering amplitudes that flow due to BCS and CDW diagrams. Several problems were proposed for further study.

Finally we discussed the possibility of singular Landau parameters and non-Fermi liquids. This was surely a possibility of strong coupling, but did not seem to happen at weak coupling within the scheme we were employing.

Although we used the modern Kadanaoff-Wilson approach to renormalization, we frequently made contact with the old field-theoretic scheme for computing the β functions. The main difference between the two schemes was that, in the former, all loop momenta were at the cutoff, while in the latter, one was at the cutoff and the rest at or below it. In Sec. II it was pointed out that in the case of the scalar field the difference did not show up in the flow of the marginal couplings at one loop, since we could set all external momenta to zero, and momentum conservation meant that, if one propagator was at the cutoff, so was the other. It was pointed out in the fermion problem that, even if we set all k's to zero, there could still be large momentum transfers of order K_F. Yet we saw repeatedly that the two approaches gave the same one-loop flow for marginal couplings. Let us recall why. First, F never flowed in either scheme, since it got its contribution at the Fermi surface and did not know about Λ in either scheme. As for V, when we set the incoming momenta on the Fermi surface at opposite angles, the total momentum vanishes, just as in the scalar field theory, and the two propagators in the loop were equal and opposite and at the cutoff in the BCS diagram. The other two diagrams did have large momentum transfers. They were suppressed in both schemes by kinematics, but not equally: the suppression was by $d\Lambda/K_F$ in the modern scheme and by Λ/K_F in the field theory scheme. This difference did not matter at the fixed point, since both factors vanished. Finally W flowed due to the ZS diagram. Even though the momentum transfer was large ($\mathbf{Q}_N$), it was such that if one propagator was at the cutoff so was the other, due to the condition $E(K)=-E(K+Q_N)$. The other two diagrams were suppressed by phase space.

The field theory scheme was the best choice for studying screening and the Kohn-Luttinger effect discussed in Appendix B.

There are many possible extensions. Inclusion of spin will produce new effects such as spin-density waves, but no new formalism is required to incorporate it. Inclusion of bosons is very interesting, but not discussed here due to lack of space. Ye and Sachdev (1991) have used the RG ideas espoused here to study a boson-fermion system describing the metal-superconductor transition. Polchinski discussed phonons in his TASI article (1992) and in a more recent preprint (1993). Tokuyasu *et al.* consider the application of the RG to finite systems like fullerenes (Tokuyasu *et al.*, 1993).

We can study disordered systems using the replica trick. Finally we can go to finite T. The fixed point discussed here will describe the crossover to $T=\infty$ in the early stages.

To conclude, the analysis in this paper has shown that the RG, which has proven its worth in critical phenomena, chaos, etc., is just as effective in helping us understand interacting fermions. Conversely, the Fermi systems, with their novel phase space and fixed-point structure, offer us a far from ordinary manifestation of the RG at work.

With various phenomena such as Fermi-liquid theory, BCS, CDW, and SDW instabilities, screening, and the Kohn-Luttinger phenomena all under the auspices of the RG, we have a better chance of solving extensions and generalizations of these problems.

ACKNOWLEDGMENTS

During the past two years I have had many occasions to discuss this material with colleagues everywhere. I have profited immensely from their generosity, interest, and criticism. It is, however, practically impossible to list all of them. As they read the paper they will no doubt find parts which they helped clarify. I must, however, acknowledge a few whose ears I bent beyond the linear regime: I. Affleck, S. Chakravarty, D. H. Lee, A. Ludwig, G. Murthy, J. Polchinski, A. Ruckenstein, H. Schulz, and D. Vollhardt. I thank E. Trubowitz for ex-

plaining some of the rigorous works to me. My thanks to my colleagues at Yale—A. Chubukov, G. Moore, J. Ye, and especially N. Read and S. Sachdev, with whom I have discussed virtually every topic on countless occasions.

This research was supported in part by a grant from the Donors of the Petroleum Research Fund and Grant DMR 9120525 from the National Science Foundation. Their support is gratefully acknowledged. I finally thank the Aspen Institute for Physics for a very stimulating summer of 1992, when many of the above-mentioned discussions took place.

APPENDIX A: COULOMB SCREENING

Here we shall use the field theory approach to study two phenomena: in this appendix the screening of the Coulomb potential and in Appendix B the Kohn-Luttinger effect. A smooth cutoff will be used on loop momenta.

We know that the instantaneous Coulomb interaction, before any mode elimination, is given by $4\pi e^2 K_F^2/Q^2$. The extra K_F^2 comes from the way our fields are normalized. The corresponding bare vertex is

$$V(4321)=4\pi e^2 K_F^2\left[\frac{1}{(K_3-K_1)^2}-\frac{1}{(K_4-K_1)^2}\right]. \tag{A1}$$

Since we shall be focusing on $K_3\simeq K_1$, we shall drop the second term.

It is generally agreed that the Coulomb potential gets screened. Should we be using a screened version here? No! Screening comes from organizing diagrammatic perturbation theory in a certain way by first summing a class of (RPA) diagrams. In the RG approach the organization is different. Since the Coulomb potential or bare coupling in the action is unambiguously known before any mode elimination, we must see what it evolves into as we carry out the RG transformation. In the field theory approach, using a smooth cutoff $e^{-\alpha|k|}$ where $\alpha=1/\Lambda$, we have to second order

$$\Gamma(Q)=V(Q\alpha)+V^2(Q\alpha)\int\frac{dk\,d\omega\,d\Omega}{(2\pi)^4}\frac{e^{-\alpha|k|}e^{-\alpha|k'|}}{[i\omega-E(K)][i\omega-E(K')]}, \tag{A2}$$

where $K'=K+Q$ and where k and $k'=|K'|-K_F$ run from $-\infty$ to ∞. We have not shown the ZS' and BCS diagrams, since they do not dominate as the ZS diagram does at small Q. Our plan is to find the β function by setting the α derivative of Γ to zero. But first let us evaluate the former more explicitly. Doing the ω integral we obtain two equal contributions from processes where a hole gets promoted to a particle and vice versa and end up with

$$\Gamma(Q)=V(Q\alpha)-V^2(Q\alpha)\int_{-\infty}^{0}\int_{-1}^{1}\frac{dk\,dz}{2\pi^2}\frac{2me^{-\alpha|k|}e^{-\alpha|k'|}\theta(k')}{Q^2+2KQz}. \tag{A3}$$

We can now do the k and z integrals. In doing so we replace K by K_F wherever appropriate and use the fact that k, k', and Q are all much smaller than K_F. For example, we set

$$k'=|K+Q|-K_F\simeq k+\frac{Q^2+2K_FQz}{2K_F} \tag{A4}$$

and so on. The result is

$$\Gamma(Q)=V-\frac{V^2m}{K_F\pi^2}\frac{1-e^{-\alpha Q}}{\alpha Q} \tag{A5}$$

$$\simeq V-\frac{V^2m}{K_F\pi^2}\frac{1}{1+\alpha Q}, \tag{A6}$$

where in the last equation we have used a simpler function with the same limits at small and large values to facilitate the analysis. The β function is now computed to be

$$\alpha\frac{dV}{d\alpha}=-\frac{V^2m}{K_F\pi^2}\frac{\alpha Q}{(1+\alpha Q)^2}. \tag{A7}$$

Notice the flow is strongest at $\alpha Q\simeq 1$, i.e., $Q\simeq\Lambda$, which makes sense in the sharp cutoff. We need a minimum momentum $Q\simeq 2\Lambda$ to scatter a particle from the shell at $-\Lambda$ to the shell at Λ. When Q is too different from this range, the flow is essentially nil. (Had we used the sharp cutoff, the β function would have had a string of θ functions, which is why we do not.) Integrating the flow from $\alpha=0$ to $\alpha=\alpha$, we obtain

$$V(Q\alpha)=\frac{1}{\frac{1}{V(Q0)}+\frac{m}{K_F\pi^2}\frac{\alpha Q}{1+\alpha Q}}, \tag{A8}$$

where $V(Q0)$ is the input potential before mode elimination. The final answer, in terms of Λ, is that

$$V(Q\alpha)=\frac{4\pi e^2K_F^2}{Q^2+\frac{Q}{Q+\Lambda}\frac{4e^2mK_F}{\pi}} \tag{A9}$$

$$\equiv\frac{4\pi e^2K_F^2}{Q^2+\frac{Q}{Q+\Lambda}\Theta^2} \tag{A10}$$

where Θ is the inverse of the Thomas-Fermi screening length.

This is the bare potential that goes into the action when the cutoff is Λ. To regain the potential before mode elimination, we must set $\Lambda=\infty$. You may ask how we can have a $\Lambda>K_F$. The point is that Λ is defined as

the inverse of the α in Eq. (A2) and not as the real cutoff for the integral. Indeed, with $\alpha=0$, the integral is still finite because it is limited by Q: for very small Q, we cannot scatter states far below K_F to states above it, and only such processes contribute to the ω integral.

Now screening refers not to the bare coupling in the action but to the full physical four-point function. It is clear, however, that at fixed Q if we send α to ∞ we kill all loops, and the bare coupling itself gives the full answer, which is

$$\Gamma(Q)=V(Q,\infty)$$

$$=\frac{4\pi e^2K_F^2}{Q^2+\dfrac{4e^2mK_F}{\pi}} . \tag{A11}$$

$$=\frac{4\pi e^2K_F^2}{Q^2+\Theta^2} . \tag{A12}$$

Let us examine what we have above. If we fix $Q\ll\Theta\ll\Theta^2/Q$ and lower Λ, this is what happens to the bare charge in the action. First, at $\Lambda=\infty$, we have the unscreened potential $4\pi e^2K_F^2/Q^2$. It has this form as long as $\Lambda\gg\Theta^2/Q$. When $Q\ll\Lambda\ll\Theta^2/Q$, it goes as $4\pi e^2K_F^2\Lambda/(\Theta^2Q)$. Finally when $\Lambda\ll Q$, we get the screened form in Eq. (A12).

We must be clear about what is done here. We tried to understand how screening takes place in the RG scheme as we eliminate modes. Since we assumed that the fermion propagator had the Fermi-liquid theory form throughout, we did not really verify that non-Fermi-liquid theory is ruled out. In other words, we must study the evolution of the fermion propagator as modified by the Coulomb potential as we go along. This was recently done using bosonization methods (Houghton *et al.*, 1993) and it was found that the static Coulomb interaction indeed gets screened.

APPENDIX B: THE KOHN-LUTTINGER EFFECT

Years ago Kohn and Luttinger pointed out that in principle any system will face the BCS instability at low temperature, even if the initial coupling is repulsive. Let us recall their argument with no reference to the RG. Consider the BCS amplitude to one loop as shown in Fig. 18:

$$\Gamma(-K_3K_3-K_1K_1)\equiv\Gamma(Q=K_1-K_3)$$

$$=V(Q)+\text{BCS}+\text{ZS}+\text{ZS}' . \tag{B1}$$

Let us compute the coefficients of the Legendre expansion

$$\Gamma_l=\int_{-1}^{1}P_l(z)\Gamma(z)dz , \tag{B2}$$

where $z=\Omega_1\cdot\Omega_3$ and l will be odd due to the Pauli principle.

The bare potential will make a contribution V_l, assumed to be positive. Since $V(z)$ is assumed to be analytic in the interval $-1\le z\le 1$, we must have $V_l\simeq e^{-l}$ in order that the infinite sum over polynomials and all derivatives converge. Now look at the one-loop corrections. They are nominally smaller, being of second order. However, their dependence on l is very interesting. This is because the ZS and ZS' graphs have singularities when $Q=K_1-K_3=2K_F$ or $Q'=K_1-K_4=2K_F$, respectively, which correspond to $z=\mp 1$. Let us focus on just ZS, since the Pauli principle will determine ZS' for us later on. Due to the singularity, the Legendre expansion coefficients fall as

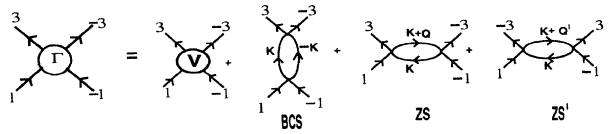

FIG. 18. The Kohn-Luttinger diagrams. Γ is the full BCS amplitude and V is the bare vertex. The ZS and ZS' diagrams have singularities when the momentum transfer equals $2K_F$, whereupon $K\simeq -1$ in ZS scatters into $K+Q\simeq 1$. Thus $V=V(\pi)$.

$$\delta\Gamma_l\simeq\frac{V_\pi^2}{l^4} , \tag{B3}$$

where $V(\pi)$, which enters both vertices, is the backward-scattering amplitude. (Let us see why. At the left vertex 1 gets scattered into 3, which equals -1, so that we have a momentum transfer $2K_F$ in the direction of 1. The loop momentum K must be nearly -1, getting knocked into $+1$ if it is to lie on or near the Fermi surface and obey momentum conservation. A similar argument applies at the other vertex.) It follows that, *if we hold V fixed and look at large l, the second term, which is attractive (for odd l, which is all we have), will dominate over the first, which is falling exponentially.* There is no question of hoping that $V(\pi)=0$, since it is (up to an overall minus sign) the sum of all the V_l's, all assumed to be non-negative.

We are trying to reproduce this in the RG language. The procedure will be just as in the screening calculation, except now we do not assume that Q is much smaller than K_F; indeed, it is nearly $2K_F$. Now we get, upon doing the ω integral and setting all unimportant factors to unity, but paying attention to the sign,

$$\Gamma(Q)=V(Q)-V^2(\pi)\int_0^\infty dk\int_{-1}^{1}dz\frac{\theta(k')e^{-\alpha k}e^{-\alpha k'}}{Q^2+2KQz} , \tag{B4}$$

where once again $\alpha=1/\Lambda$, and the primed quantities refer to $K'=K+Q$. Let us define

$$x=2K_F-Q . \tag{B5}$$

By drawing a sketch of the Fermi surface you may verify that since all the action is near $Q=2K_F$,

$$e^{-\alpha k'}\simeq e^{-\alpha(k-x)} , \tag{B6}$$

and that $-1 \le z \le 1$ for $k > x$, while $z_m \le z \le 1$ for $k < x$, with z_m being the point where $k' = 0$. Putting all this into Eq. (B4), we obtain

$$\Gamma = V - V^2(\pi) \left[e^{\alpha x} \int_x^\infty \frac{dx}{2KQ} \ln \left[\frac{Q+2K}{Q-2K} \right] e^{-2\alpha k} + e^{\alpha x} \int_0^x \frac{dk}{2KQ} \ln \left[\frac{Q^2+2KQ}{2K_F k} \right] e^{-2\alpha k} \right] . \tag{B7}$$

By subtracting the second integral from 0 to ∞, which does not affect the singularity in question, and shifting the origin and rescaling k, we get

$$\Gamma = V - V^2(\pi) e^{-\alpha x} \int_0^\infty e^{-\alpha y} \ln \left[\frac{y+2x}{y+x} \right] dy . \tag{B8}$$

We shall drop the $e^{-\alpha x}$, since it does not modify the singularity at $x = 0$. Next we approximate as follows:

$$\begin{aligned} x &= 2K_F - Q \\ &= 2K_F(1 - \sin\theta_{13}/2) \\ &\simeq K_F(1 - \sin\theta_{13}/2)(1 + \sin\theta_{13}/2) \\ &= K_F(1+z) . \end{aligned} \tag{B9}$$

Hereafter we shall use $K_F = 1$. We now do an integration by parts, throw out the surface term, and obtain

$$\Gamma = V - \frac{V^2(\pi)}{\alpha} \int_0^\infty e^{-\alpha y} \left[\frac{1}{y+1+z} - \frac{2}{2y+1+z} \right] dy . \tag{B10}$$

Now do the angular momentum transform using

$$Q_l(z_0) = \frac{1}{2} \int_{-1}^{1} \frac{P_l(z) dz}{z_0 - z} \tag{B11}$$

to obtain

$$\Gamma = V + \frac{V^2(\pi)}{\alpha} \int_1^\infty Q_l(y) [e^{\alpha/2} e^{-\alpha y/2} - e^{\alpha} e^{-\alpha y}] dy . \tag{B12}$$

Next we use the result that as $l \to \infty$,

$$Q_l(y) \to \sqrt{2/\pi X}\, e^{-X} , \tag{B13}$$

$$X = l\sqrt{y^2 - 1} , \tag{B14}$$

to obtain for large l and α

$$\Gamma = V + \frac{V^2(\pi)}{\alpha l^2} \int_0^\infty e^{-X} \sqrt{X} \, [e^{-\alpha X^2/2l^2} - e^{-\alpha X^2/4l^2}] dX \tag{B15}$$

$$\equiv V + L . \tag{B16}$$

Note that

$$L(\alpha/l^2 \to \infty) \simeq -\frac{c}{l^{1/2}\alpha^{7/4}} , \tag{B17}$$

$$L(l^2/\alpha \to \infty) \simeq \frac{-c'}{l^4} . \tag{B18}$$

We fit L with a simpler function with the same limits:

$$\Gamma = V - \frac{V^2(\pi)}{\alpha l^2} R(\alpha/l^2) , \tag{B19}$$

$$R(x) = \frac{x}{1 + x^{7/4}} . \tag{B20}$$

If we calculate the β function (including a factor of 2 due to the ZS′ diagram) and also include the usual contribution from the BCS diagram, we get the result quoted in the text:

$$\frac{dV_l}{dt} = -\frac{V_l^2}{4\pi} - \frac{V^2(\pi)\lambda^{7/4}}{l^{15/2}[\lambda^{7/4} + l^{-7/2}]^2} . \tag{B21}$$

REFERENCES

Abrikosov, A. A., L. P. Gorkov, and I. E. Dzyaloshinski, 1963, *Methods of Quantum Field Theory in Statistical Mechanics* (Dover, New York).

Anderson, P. W., 1984, *Basic Notions of Condensed Matter Physics* (Benjamin-Cummings, Menlo Park).

Anderson, P. W., 1990, Phys. Rev. Lett. **64**, 1839.

Anderson, P. W., and G. Yuval, 1970, Phys. Rev. Lett. **23**, 89.

Baranov, M. A., A. V. Chubukov, and M. Kagan, 1992, Int. J. Mod. Phys. **6**, 2471.

Bares, P., and X. G. Wen, 1993, Phys. Rev. B **48**, 8636.

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

Benfatto, G., and G. Gallavotti, 1990, Phys. Rev. B **42**, 9967.

Berezin, F. A., 1966, *The Method of Second Quantization* (Academic, New York).

Boninsegni, M., and E. Manousakis, 1991, Phys. Rev. B **43**, 10353.

Bourbonnais, C., and L. G. Caron, 1991, Int. J. Mod. Phys. B **5**, 1033.

Brazovskii, S. A., 1975, Zh. Eksp. Teor. Fiz. **68**, 175 [Sov. Phys. JETP **41**, 85 (1975)].

Callan, C. G., 1970, Phys. Rev. D **2**, 1541.

Castellani, C., C. Di Castro, and W. Metzner, 1993, "Dimensional Crossover from Fermi to Luttinger Liquid," preprint No. 954, Rome.

Castro Neto, A. H., and E. Fradkin, 1993, "Bozonization of the Low Energy Excitations of Fermi Liquids," University of Illinois preprint, 1993.

Di Castro, C., and G. Jona Lasinio, 1969, Phys. Lett. A **29**, 322.

Di Castro, C., and W. Metzner, 1991, Phys. Rev. Lett. **67**, 3852.

Di Castro, C., and W. Metzner, 1992, Phys. Rev. Lett. **69**, 1703.

Elser, V., D. Huse, B. Shraiman, and E. D. Siggia, 1990, Phys. Rev. B **41**, 6715.

Engelbrecht, E., and M. Randeria, 1990, Phys. Rev. Lett. **65**, 1032.

Feldman, J., J. Magnen, V. Rivasseau, and E. Trubowitz, 1992, Helv. Phys. Acta **65**, 679.

Feldman, J., J. Magnen, V. Rivasseau, and E. Trubowitz, 1993, Helv. Phys. Acta **66**, 498. See also "An intrinsic $1/N$ expan-

sion for many-fermion systems," Europhys. Lett. (to be published).
Feldman, J., and E. Trubowitz, 1990, Helv. Phys. Acta **63**, 157.
Feldman, J., and E. Trubowitz, 1991, Helv. Phys. Acta **64**, 213.
Fisher, M. E., 1974, Rev. Mod. Phys. **4**, 597.
Fisher, M. E., 1983, in *Critical Phenomena,* edited by F. W. J. Hahne, Lecture Notes in Physics No. 186 (Springer, Berlin), pp. 1–139.
Gell-Mann, M., and F. E. Low, 1954, Phys. Rev. **95**, 1300.
Goldenfeld, N., 1992, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Reading, MA).
Haldane, F. D. M., 1981, J. Phys. C **14**, 2585.
Hertz, J. A., 1976, Phys. Rev. B **14**, 1165.
Houghton, A., and J. B. Marston, 1993, Phys. Rev. B **48**, 7790.
Houghton, A., H. J. Kwon, J. B. Marston, and R. Shankar, 1993, "Coulomb interaction is screened in bosonized Fermi liquids," Brown University Preprint.
Ioffe, L. B., and G. Kotliar, 1990, Phys. Rev. B **42**, 10348.
Itzykson, C., and J. B. Zuber, 1980, *Quantum Field Theory* (McGraw Hill, New York).
Itzykson, C., and J.-M. Drouffe, 1989, *Statistical Field Theory* Vol. I (Cambridge University Press, Cambridge, England).
Kadanoff, L. P., 1965, Physica **2**, 263.
Kadanoff, L. P., 1977, Rev. Mod. Phys. **49**, 267.
Kogut, J. B., 1979, Rev. Mod. Phys. **51**, 659.
Kohn, W., and J. Luttinger, 1965, Phys. Rev. Lett. **15**, 524.
Krishna-murthy, H. R., J. W. Wilkins, and K. G. Wilson, 1980, Phys. Rev. B **21**, 1003, 1044.
Landau, L. D., 1956, Zh. Eksp. Teor. Fiz. **30**, 1058 [Sov. Phys. JETP **3**, 920].
Landau, L. D., 1957, Zh. Eksp. Teor. Fiz. **35**, 97 [Sov. Phys. JETP **8**, 70].
Landau, L. D., 1959, Zh. Eksp. Teor. Fiz. **32**, 59 [Sov. Phys. JETP **5**, 101].
Le Bellac, M., 1991, *Quantum and Statistical Field Theory* (Oxford University Press, New York).
Lee, P. A., and N. Nagaosa, 1990, Phys. Rev. Lett. **64**, 2450.
Lee, P. A., and T. V. Ramakrishnan, 1985, Rev. Mod. Phys. **57**, 287.
Leggett, A. J., 1975, Rev. Mod. Phys. **47**, 331.
Lifshitz, E. M., and L. P. Pitayevskii, 1980, *Statistical Physics II* (Pergamon, Oxford).
Ludwig, A. W. W., and I. Affleck, 1991a, Phys. Rev. Lett. **67**, 3160.
Ludwig, A. W. W., and I. Affleck, 1991b, Nucl. Phys. B **360**, FS, 641.
Ludwig, A. W. W., and I. Affleck, 1992, Phys. Rev. Lett. **68**, 1046.
Luttinger, J. M., 1960, Phys. Rev. **119**, 1153.
Luttinger, J. M., 1961, Phys. Rev. **121**, 942.
Ma, S.-K., 1976, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA).
Mahan, G. D., 1981, *Many-body Physics* (Plenum, New York).
Mermin, D., 1967, Phys. Rev. **159**, 161.
Metzner, W., C. Di Castro, 1993, Phys. Rev. B **47**, 16, 107.
Millis, A., 1993, Phys. Rev. B **48**, 7183.
Morel, P., and P. W. Anderson, 1962, Phys. Rev. **125**, 1263.
Negele, J. F., and H. Orland, 1988, *Quantum Many-Particle Systems* (Addison Wesley, New York).
Nozières, P., 1964, *Interacting Fermi Systems* (Benjamin, New York).
Nozières, P., 1974, J. Low. Temp. Phys. **17**, 31.
Perakis, I., C. M. Varma, and A. E. Ruckenstein, 1993, Phys. Rev. Lett. **70**, 3467.
Pines, D., and P. Nozières, 1966, *The Theory of Quantum Liquids* (Addison-Wesley, Menlo Park).
Plischke, M., and B. Bergerson, 1989, *Equilibrium Statistical Physics* (Prentice Hall, Englewood Cliffs, New Jersey).
Polchinski, J., 1992, "Effective Field Theory and the Fermi Surface," Proceedings of 1992 Theoretical Advanced Studies Institute in Elementary Particle Physics, edited by J. Harvey and J. Polchinski (World Scientific, Singapore, 1993).
Polchinski, J., 1993, "Low Energy Dynamics of the Spinon-Gauge System," preprint NSF-ITP-93-93, UTTG-09-93.
Sachdev, S., 1989, Phys. Rev. B **39**, 12232.
Sachdev, S., 1993, "Quantum Phase Transitions and Conserved Charges," Condens. Matter 9312018.
Schulz, H. J., 1991, Int. J. Mod. Phys. B **43**, 10353.
Schwinger, J., 1970, *Quantum Kinematics and Dynamics* (Benjamin, New York).
Shankar, R., 1991, Physica A **177**, 530.
Solyom, J., 1979, Adv. Phys. **28**, 201.
Stamp, P. C. E., 1992, Phys. Rev. Lett. **68**, 2180.
Stamp, P. C. E., 1993, J. Phys. (Paris) **3**, 625.
Stueckelberg, E. C. G., and A. Petermann, 1953, Helv. Phys. Acta **26**, 499.
Swift, J., and P. C. Hohenberg, 1977, Phys. Rev. A **15**, 319.
Symanzik, K., 1970, Commun. Math. Phys. **18**, 227.
Tokuyasu, T., M. Kamal, and G. Murthy, 1993, "Numerical Renormalization Group for Finite Fermi Systems," Phys. Rev. Lett. (to be published).
Trugman, S., 1988, Phys. Rev. B **37**, 1597.
Varma, C. M., P. Littewood, S. Schmitt-Rink, E. Abrahams, and A. E. Ruckenstein, 1989, Phys. Rev. Lett. **63**, 1996.
Vollhardt, D., and P. Wölfle, 1990, *The Superfluid Phases of Helium 3* (Taylor & Francis, London).
Weinberg, S., 1986, Prog. Theor. Phys. (Supplement) **86**, 43.
Weinberg, S., 1993, "Effective Action and Renormalization Group Flow of Anisotropic Superconductors," University of Texas preprint UTTG-18-93.
Wen, X. G., 1990, Phys. Rev. B **42**, 6623.
Wilson, K. G., 1971, Phys. Rev. B **4**, 3174 (Part I), 3184 (Part II).
Wilson, K. G., 1975, Rev. Mod. Phys. **47**, 773.
Wilson, K. G., and M. E. Fisher, 1977, Phys. Rev. Lett. **28**, 240.
Wilson, K. G., and J. B. Kogut, 1974, Phys. Rep. **12**, 7.
Yang, C. N., and C. P. Yang, 1976, Phys. Rev. **150**, 321.
Ye, J., and S. Sachdev, 1991, Phys. Rev. **44**, 10173.
Zimanyi, G., and K. Bedel, 1991, Phys. Rev. Lett. **66**, 228.
Zinn-Justin, J., 1989, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford).