6

# Theory of the One-Dimensional Electron Gas

*V. J. Emery*

## 1. Basic Physics

### 1.1. Introduction

The properties of organic metals and other highly conducting solids are far too complex to be understood by means of purely microscopic calculations. They involve an intricate interplay of electronic, molecular, and lattice motion, modified by disorder and conditioned by the symmetry of the underlying structure, all of which must be included in some measure if all of the experiments are to be understood. At the same time the chainlike structure of the lattice often confines the motion of the electrons to one direction in space, and this results in an exaggerated role of the interactions as well as subtle cancellation between various many-body processes which are simply not revealed by the approximations usually used to obtain at least a qualitative description of more isotropic systems. Intuition developed in this way cannot be trusted.

Faced with this situation, it is fortunate that the restriction to one-dimensional motion brings along a compensating possibility of finding exact solutions of certain many-body problems. They are not necessarily the precise models we want to study for real physical systems, but they do frequently contain so much of the essence that the solutions are able to offer some guidance about what is going on as well as to show where more approximate theories are to be trusted. Often no more is needed, for truly microscopic calculations rarely confront experiment: it is much more usual to get a feeling for what qualitative features can appear, to make limited

*V. J. Emery* • Department of Physics, Brookhaven National Laboratory, Upton, New York 11973.

predictions about the dependence of physical properties on temperature, frequency, etc., then to introduce phenomenological parameters which are fitted to experiment. Landau's Fermi liquid theory[1] is a prime example of this procedure, and it has been employed to great advantage in understanding the properties of liquid $^3$He, in which strong-coupling effects are so very important.

It is useful to pursue a similar course in the study of organic metals, and this chapter is concerned with the first step—a discussion of exact solutions or controlled approximations of the properties of an electron gas in one dimension. This is a limited goal. For the most part, the effects of phonons and molecular vibrations will be ignored and there will be little consideration of coupling between chains or attempt to use the results to parametrize the experiments. The latter in particular involves a certain amount of taste in adapting the results to the physical systems under consideration and it is best considered case by case. Some of these questions are addressed in other chapters of this volume.

Three main approaches have been widely used in this field: Bethe's ansatz,[2] the renormalization group method, and boson representations. The first has produced many important results, some of which will be used later, but it does not tell us as much as we should like to know about the correlation functions which are so important for revealing the nature of the collective motion or the possible phase transitions which the system may undergo. On the other hand the other two methods do this very nicely even though they omit almost all effects of the lattice. Despite the importance of Bethe's ansatz, this chapter will concentrate largely upon the other two methods, for they embody the most recent technical developments and also have a bearing on problems of high-energy physics.

The boson representations are actually much simpler to use than many of the more common methods of many-body theory and, for this reason, a good part of Section 2 is devoted to a rather detailed account of the technique so as to give the interested reader the necessary background to follow the subsequent discussion and to carry out similar calculations for himself. Before plunging into the details, however, it is perhaps useful to give a more qualitative survey of some of the physical questions to be addressed.‡

### 1.2. Phase Transitions and Long-Range Order

Much of the physics of almost one-dimensional metals is concerned with understanding the conditions for various kinds of phase transitions that lead to collective states governing the low-temperature behavior. This

‡ The discussion given in the rest of this section has been taken from the author's lectures[3] at the NATO school on the Chemistry and Physics of One-Dimensional Metals held in Bolzano, Italy, August, 1976.



is of central importance in the search for high-temperature superconductivity, which, for example, could be prevented by the intervention of a metal-insulator transition. The most frequently discussed transitions are to states showing singlet superconductivity, triplet superconductivity, static charge-density waves (CDW), or static spin-density waves (SDW). Superconductivity requires the formation of bound pairs of electrons which are (roughly speaking) bosons and form a superfluid. The pairs may form with singlet spin and even angular momentum as in a metal[4] or with a triplet spin and odd angular momentum[5] as in liquid $^3$He (the latter is a superfluid but not a superconductor because $^3$He atoms are not charged). In a static CDW state, the electron density is not uniform but has a periodic modulation, which, in the simplest situation, has the form

$$n(\mathbf{r}) = n_0 + n_1 \cos(\mathbf{q}\cdot\mathbf{r} + \phi) \tag{1}$$

at position $\mathbf{r}$. Such states have been studied extensively in layered compounds. In an SDW state it is the spin density that varies from point to point to give a loose kind of antiferromagnetism. All of these transitions have been observed in more isotropic materials, but there may be others, as yet undiscovered, that are peculiar to almost one-dimensional metals.

The low-temperature phases are characterized by the existence of a new kind of order which extends throughout the system. They are described by one or more *order parameters*—in Equation (1) by $n_1$ and $\phi$, the amplitude and phase of the CDW. The magnitude of $n_1$ is determined by minimizing the free energy, but unless there is a specific coupling to something external (such as the lattice), $\phi$ is free to vary and the state is degenerate. Above the transition temperature, particularly in a nearly one-dimensional metal, the order can extend over quite long but finite distances. Within such regions the system has the appearance of the ordered phase. As the temperature is lowered, the range of the order increases and it becomes infinite at the transition temperature $T_c$.

However, in a purely one-dimensional system with short-range forces, thermal or quantum fluctuations prevent the formation of long-range order and there is no phase transition at all. To see how this comes about, imagine that the system breaks into two segments with the phase of the order parameter taking one value to the left of a point $P$ and another to the right. There are two contributions to the change $\delta F$ in the free energy—the surface free energy $\sigma$ due to the mismatch at $P$ and an entropy $\ln M$ which arises because $P$ can be at any of the $M$ sites:

$$\delta F = \sigma - T \ln M \tag{2}$$

Since the contact between segments occurs at a point, $\sigma$ is finite for large $M$ and the gain from the entropy far outweighs the cost in surface free energy, so $\delta F < 0$. Then it is favorable to break the system into further segments

and to continue to to do this until there is no remnant of the original long-range order. This argument does not work in two (or more) dimensions because, if the region of fluctuation contains a macroscopic number $\lambda M$ of particles, the surface energy $\sigma$ is proportional to $M^{1/2}$ [or to $M^{1-1/d}$ in $d$ dimensions] and is bigger than $\ln M$. The existence of a fluctuation then depends upon the sign of the surface term. The discussion is also modified if there are long-range forces which give interactions between the segments. But none of this concerns us here; the point is that fluctuations must be taken seriously and it is not useful to do a rough calculation of the properties of a one-dimensional system. This is true even at $T=0$, where the entropy contribution to Equation (2) vanishes, because, in general, the transition is removed by quantum fluctuations which, so far, have been ignored.

Even though there is no long-range order, it is possible for very-long-range (but finite) correlations to build up, and they trigger a phase transition in real materials by enhancing the coupling between the chains. The important question then becomes the determination of what kind of correlations grow the most rapidly as the temperature is lowered, for they, in conjunction with the details of the interchain coupling, control the nature of the phase transitions that finally take place.

There is a further effect severely limiting the applicability of simple theories. At low temperatures, the various kinds of phase transitions are consequences of interactions between particles primarily on the Fermi surface, but in one dimension, the Fermi "surface" consists of the two points $k=\pm k_F$, where $k_F$ is the Fermi wave vector. Then the available phase space is so small that the various kinds of singular process cannot help but interfere with each other, so it is not useful to do a theory of one kind of transition and to include fluctuations in a self-consistent way—the situation virtually demands an exact solution, or at least controlled approximations which can perhaps be obtained in certain limiting cases. These are the questions with which we shall be concerned.

### 1.3. Mathematical Model

The discussion will be based upon the Hamiltonian

$$H = H_0 + H_1 + H_2 \tag{3}$$

where

$$H_0 = -\varepsilon \sum_{n,\sigma} (\psi^\dagger_{n\sigma}\psi_{n+1,\sigma} + \psi^\dagger_{n+1,\sigma}\psi_{n\sigma}) \tag{4}$$

Hopping

$$H_1 = \frac{U}{2}\sum_{n,\sigma} \rho_{n\sigma}\rho_{n,-\sigma} \tag{5}$$

Hubbard



and

$$H_2 = V \sum_{m,\sigma,\sigma'} \rho_{m\sigma}\rho_{m+1,\sigma'} \tag{6}$$

"extended" density density

where $\psi^\dagger_{n,\sigma}$ creates a Fermion of spin $\sigma$ (of value $\uparrow$ or $\downarrow$) at a lattice site $n$ and $\rho_{m\sigma} = \psi^\dagger_{m\sigma}\psi_{m\sigma}$ is the number operator. This Hamiltonian describes what is often called an extended Hubbard model, with $H_0$ producing hopping from site to site and $H_1$ and $H_2$ giving an on-site or an intersite interaction. The couplings $U$ and $V$ come partly from Coulomb interactions and partly from exchange of phonons or molecular vibrations, but, for the most part, retardation will be ignored and the interactions will be of short range in order to focus on the Fermi-surface effects, which are so important in one dimension. Even with these restrictions there is a rich variety of physical phenomena, and the theory forms a good basis for extension to more complicated situations. Later, in discussing the continuum limit, it will prove useful to generalize $H$ by allowing a different coupling constant for different scattering processes, but Equations (3)–(6) will always be kept in mind as the more physical model that has to be solved.

There are certain kinematic properties of the noninteracting system $H_1 = H_2 = 0$ which are of importance in interpreting the physical effects. Imagine that there are $M$ molecular sites in a length $L \equiv Ms$, beyond which the system is extended periodically. It is assumed in Equations (3)–(6) that each site has just one spatial state for an electron to occupy. This is an oversimplification, but it seems to be a good approximation for many physical systems.[6] The free Hamiltonian $H_0$ may be diagonalized by Fourier transformation

$$\psi^\dagger_{m\sigma} = M^{-1/2}\sum_k e^{ikms} a^\dagger_{k\sigma} \tag{7}$$

where the wave vectors $k$ have values $2\pi\nu/L$ with $\nu$ an integer. The energy spectrum is $(-2\varepsilon \cos ks)$ and the ground state is a Fermi sea with two electrons in each state of wave vector $k$ less than the Fermi wave vector $k_F \equiv 2\pi N_0/L$. The value of $k_F$ is related to the number $N_e$ of electrons by

$$\begin{aligned} N_e &= \sum_{n=-N_0}^{N_0} 2 \\ &= 4N_0 \\ &= 2k_F L/\pi \end{aligned} \tag{8}$$

which may be expressed in terms of the reciprocal lattice vector $G \equiv 2\pi/s$ as

$$N_e/M = 4k_F/G \tag{9}$$

This relationship may be used to calculate $N_e$ once $k_F$ and $G$ are obtained from x-ray and neutron scattering experiments. An important special case is a half-filled band $N_e = M$ for which $G = 4k_F$.

### 1.4. Strong Coupling

A particularly simple picture of the possible states of the system may be obtained in the large-$|U|$ limit, in which $H_0$ and $H_2$ are treated as perturbations. It is not necessarily assumed that large $|U|$ corresponds to any particular material, and for the present purpose it is merely a limit in which the properties of one-dimensional conductors can easily be visualized. The discussion will largely be qualitative at this stage and the mathematical basis will be given in Section 3. The properties depend upon the sign of $U$.

#### 1.4.1. $U<0$

When $H_0 = 0 = H_2$, the electrons occupy molecular sites in pairs to take advantage of the on-site attraction. This is illustrated in Figure 1a, where the crosses represent molecules and the arrows refer to up or down spin electrons. The ground state is very degenerate because the energy does not depend upon which sites are occupied.

*CDW states* occur in an extreme form when there is an intersite repulsion $V_{ij}$ but still no hopping. To minimize the energy, the pairs are equally spaced, as shown in Figure 1b for a half-filled band. The charge

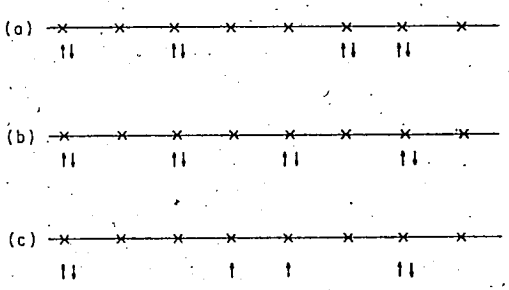


Figure 1. Some configurations of electrons for the strong coupling limit with $U<0$. The arrows show the electron spin.



density varies periodically from 1 to 0 in a distance $L/N_e$ so, using Equation (8), the wave vector is just $2k_F$. This is a possible ground state because the system is classical when $H_0 = 0$. More realistically, when $H_0 \neq 0$, the picture is not so static and there is a much smaller modulation of the charge density. Also, if $N_e/M$ is irrational, the electrons cannot be distributed so neatly amongst the sites. However, it will be seen that the wave vector $2k_F$ always characterizes the CDW.

*Singlet superconductivity* can arise when hopping ($H_0$) is included. The electron pairs are bosons, bound in a singlet state, and it is possible that they become superfluid (and hence superconducting since they are charged) at low enough temperatures. Actually this does not happen in a purely one-dimensional system since quantum mechanical fluctuations prevent superconductivity even at zero temperature. When hopping between chains is allowed, a transition does take place at a temperature $T_c$ determined partly by the value of the perpendicular hopping amplitude. In general, $T_c$ is not the same as the temperature $|U|/k_B$, which characterizes pair formation, a distinction which persists for weak coupling. In this respect nearly one-dimensional conductors differ from isotropic three-dimensional metals, where, in weak coupling, the pairs form and Bose-condense at the same temperature.[4]

*Triplet superconductivity* will not occur because the electrons are bound into singlet pairs before long-range triplet correlations can build up.

*The excited states* are of two kinds. Charge-density wave excitations require the movement of pairs from site to site and they can be phonons or plasmons according to whether the neutralizing background (ions or holes) moves in phase or out of phase with the electrons. However, in a spin-wave excitation, a spin is turned over and a pair must be broken since two electrons of the same spin cannot occupy the same site. This costs an energy $|U|$, which appears as a *gap in the spin-wave spectrum*. A static version of such an excitation is shown in Figure 1c. In the same way, the Pauli susceptibility $\chi$ is proportional to $e^{-|U|/k_BT}$ at low temperatures since only thermally broken pairs can respond to a weak magnetic field. The spin-wave gap and exponential susceptibility are general features of one-dimensional systems with attractive interactions and they occur for weak coupling also. Later it will be shown that, in the field theory models, they are related to the excitation of soliton–antisoliton pairs.

#### 1.4.2. $U>0$

In this case the electrons avoid the strong repulsive on-site interaction and the ground state has no doubly occupied sites. (The discussion will assume $N_e \leq M$ but it is easily adapted to $N_e > M$.) When $H_1 = 0 = H_2$,

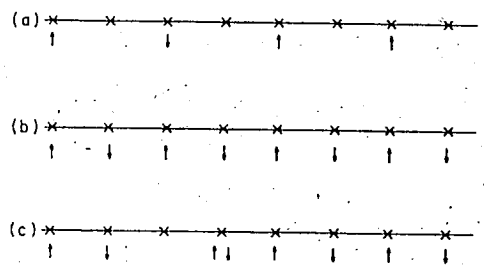


Figure 2. some static configurations for the strong coupling limit with $U > 0$. The arrows show the electron spin.

there is degeneracy because neither the spin directions nor the locations of the unoccupied sites have any effect on the energy.

*CDW states* occur again in an extreme form when there is an intersite repulsion $V_{mn}$ but not hopping. In contrast to $U<0$, however, single electrons rather than pairs are equally spaced so the period of the CDW is halved and its wave vector is $4k_F$. This is shown in Figure 2a for a quarter-filled band, which is very close to the situation in TTF-TCNQ. (No particular significance is attached to the spin orientations at this stage.) Once again, hopping makes the CDW weaker and less static. It also mixes in doubly occupied sites and restores the Fermi sea, which may lead to an additional $2k_F$ periodicity by the mechanism described for $U<0$.

*The SDW instability* is most clearly visualized for a half-filled band, which has exactly one electron per site so that only the spin degrees of freedom have to be considered. Virtual hopping produces an effective antiferromagnetic exchange interaction and the ground state has a modulation of the spin density, which is illustrated in Figure 2b. (This picture is correct if only the $z$ components of the spins are coupled.) The period is $2L/N_e$ and the wave vector is $2k_F$. For weaker coupling and a different number of electrons, the state is more dynamic, but the wave vector is still $2k_F$.

*Superconductivity* will occur when $N_e \neq M$ and $V$ is attractive. The electrons form Cooper pairs[4], with either singlet[4] or triplet[5] spin. It requires a coupling to the lattice or to collective modes of the electrons themselves[7] to overcome the intersite Coulomb interactions and make $V_{mn}$ attractive. As for $U<0$, hopping between the chains is required for the phase transitions to take place.

*The excited states* are rather different in character from those described for $U<0$. When $N_e \neq M$, there is no difficulty in rearranging spins and charges to produce an excited state but, *for a half-filled band, the CDW excitations have an energy gap*, for they require the double occupancy of a site as shown in Figure 2c. The spin waves are known exactly for this case and have no gap.



It is very useful to keep this overall picture in mind when following the mathematical description of one-dimensional conductors since it gives a concrete feeling for what is going on, even though, for small $|U|$, the states are not so sharply delineated.

## 2. Spinless Fermions

### 2.1. Definition of the Continuum Limit

s is Lattice constant.

This section will be concerned with the theory of spinless fermions in one dimension for which it is possible to find an exact solution in the continuum limit $s \to 0$. It is interesting to consider this limit because distances are given by $ps$ with $p \to \infty$ as $s \to 0$ and thus the correlation functions reflect the long-range behavior (large $p$) of the lattice system, and this is just what is required for a discussion of long-range order.

The starting point is the lattice model introduced in Section 1.3 and, initially, the continuum limit will be described for the free fermion Hamiltonian

$$H_0 = -\varepsilon \sum_n (\psi_n^\dagger \psi_{n+1} + \psi_{n+1}^\dagger \psi_n) \tag{10}$$

which is the same as $H_0$ in Equation (4) except that the spin label has been omitted. The fermion creation and annihilation operators obey the anticommutation relations

$$[\psi_m, \psi_n^\dagger]_+ = \delta_{mn} \tag{11}$$

and in the Heisenberg picture (with $\hbar = 1$), the equations of motion are

$$i\frac{\partial \psi_n(t)}{\partial t} = [\psi_n(t), H]$$

$$= \varepsilon[\psi_{n+1}(t) + \psi_{n-1}(t)] \quad \checkmark \tag{12}$$

which can readily be solved by Fourier transformation, using Equation (7), to obtain

$$a_k(t) = e^{2i\varepsilon t \cos ks} a_k(0) \quad \checkmark \tag{13}$$

Thus the time evolution is determined by the cosine spectrum of the discrete lattice.

In taking the continuum limit, it is necessary to take account of the fact that, in Equation (12), the time derivative of fields on odd (even) sites is related to the values of the fields themselves on even (odd) sites. This property may be preserved by introducing two functions $\psi_E(x, t)$ and $\psi_0(x, t)$ such that, as $s \to 0$,

$$\psi_{2n}(t) \to i^{2n}(2s)^{1/2}\psi_E(x, t)$$
$$\psi_{2m+1}(t) \to i^{2m+1}(2s)^{1/2}\psi_0(y, t) \tag{14}$$

where $n$ and $m \to \infty$ as $s \to 0$ in such a way that $2ns \to x$ and $(2m+1)s \to y$. The factors $s^{1/2}$ are introduced to turn Equations (11) into continuum anticommutation relations with $\delta_{mn} \to \delta(x-x')$, and the phase factors change $(\psi_{n+1}+\psi_{n-1})$ into $i^{n+1}(\psi_{n+1}-\psi_{n-1})$ which will become a derivative in the continuum limit. Then if $2\varepsilon s$ tends to a finite limit $v_F$, Equation (12) becomes.

$$\lim_{s \to 0} 2\varepsilon s = v_F$$

$$\frac{\partial \psi_E}{\partial t} = v_F \frac{\partial \psi_0}{\partial x} \tag{15}$$

$$\frac{\partial \psi_0}{\partial t} = v_F \frac{\partial \psi_E}{\partial x} \tag{16}$$

which is the two-component Dirac equation for massless particles. The velocity $v_F$ will play the role of the Fermi velocity for the one-dimensional electron gas, but it is analogous to the velocity of light in the Dirac equation and, in this sense, the whole theory will turn out to be Lorentz invariant. Equations (15) and (16) may be separated by defining

$$\bar{\psi}_1 = (\psi_E + \psi_0)/2^{1/2} \tag{17}$$
$$\bar{\psi}_2 = (\psi_E - \psi_0)/2^{1/2} \tag{18}$$

so that

$$\frac{\partial \bar{\psi}_1}{\partial t} = v_F \frac{\partial \bar{\psi}_1}{\partial x} \tag{19}$$

$$\frac{\partial \bar{\psi}_2}{\partial t} = -v_F \frac{\partial \bar{\psi}_2}{\partial x} \tag{20}$$

The new fields $\bar{\psi}_1$ and $\bar{\psi}_2$ define fermions moving to the right ($\bar{\psi}_1$) or left ($\bar{\psi}_2$) with velocity $v_F$. The energy spectra are $\pm v_F k$ instead of $2\varepsilon \cos ks$ as found in the lattice model, and there is no lower cutoff so the energy of a particle becomes $-\infty$ when $k \to \mp\infty$. This implies that there is no ground state and, as usual in Dirac theory, it requires the vacuum to have all



negative energy states filled. This method of going between lattice and continuum forms of the Dirac equation was introduced by Kogut and Susskind.[8] It should be mentioned that it is not obvious that taking the limit $s \to 0$ in the equations of motion will give the same results as taking that limit in correlation functions and, indeed, there are cases where it is not true. For the present purpose Equations (19) and (20) *define* the free fermion part of the model and then the question of interchange of limits will not arise.

For a many-particle problem it is necessary to fill positive energy states up to $k = k_F$ for right-going particles and $k = -k_F$ for left-going particles. Then it is convenient to introduce new fields

$$\psi_1 = e^{ik_F(v_F t - x)}\bar{\psi}_1 \tag{21}$$

$$\psi_2 = e^{ik_F(v_F t + x)}\bar{\psi}_2 \tag{22}$$

which satisfy Equations (19) and (20) and the anticommutation relations, so that the phase factors explicitly eliminated from $\psi_1$ and $\psi_2$ will no longer appear in the manipulations of operators. This simplifies the discussion. The presence of the time-dependent factors in Equations (21) and (22) implies that $v_F k_F(\hat{N}_1 + \hat{N}_2)$ has been subtracted from $H_0$ ($\hat{N}_1$ and $\hat{N}_2$ are number operators for right- and left-going particles) and its continuum form is

$$H_0 = iv_F \int_0^L dx \left[\psi_1^\dagger(x)\frac{\partial}{\partial x}\psi_1(x) - \psi_2^\dagger(x)\frac{\partial}{\partial x}\psi_2(x)\right] \tag{23}$$

In the absence of spin, the interaction $H_1$ does not occur, and introducing the definitions of $H_1$ and $H_2$ into Equation (6) gives the continuum limit in the form

$$H_2 = \int dx\, [V_1 s(\rho_1^2 + \rho_2^2) + 2V_2 s\rho_1\rho_2] \tag{24}$$

apart from a multiple of the number operator. In Equation (24), $\rho_i = \psi_i^\dagger(x)\psi_i(x)$ and the couplings $V_1 = V$ and $V_2 = 2V$ have been labeled so that their different roles in the calculation may be distinguished. The continuum limit for spinless fermions given by the Hamiltonian $H_0 + H_2$ is essentially the model proposed by Luttinger[9] for a one-dimensional system, and it was shown by Mattis and Lieb[10] that the partition function can be evaluated exactly. Their method is the first step in introducing boson representations of Fermion operators.

### 2.2. Boson Representations and the Free Energy

In general, it is expected that density operators commute with each other because they are functions of the coordinates of the particles and not their momenta. However, in a second-quantized theory, the commutation relations may depend upon the form of the vacuum and they can be different if the negative energy states are occupied or unoccupied. To see how this comes about, consider first the right-going branch for which the Fourier-transformed annihilation operators are defined by the continuum version

$$a_{1k} = L^{-1/2} \int dx\, e^{ikx} \psi_1(x) \tag{25}$$

with the aid of Equation (14). They satisfy the commutation relations

$$[a_{1k}, a^\dagger_{1k'}]_+ = \delta_{kk'} \tag{26}$$

The canonical transformation to operators with a vacuum containing filled negative energy states is given by

$$\begin{aligned} \bar{a}_{1k} &= a^\dagger_{1k}, \qquad k<0 \\ &= a_{1k}, \qquad k>0 \end{aligned} \tag{27}$$

Suppose that the momenta $k$ lie in a range $-K \le k \le K$ where, eventually, $K$ will become infinite. Then the Fourier transform $\int dx\, \rho_1(x) \exp(ipx)$ becomes

$$\begin{aligned} \rho_{1p} &= \sum_{k=-K}^{K-p} a^\dagger_{1,k+p} a_{1k} \\ \rho_{1,-p} &= \sum_{k=-K}^{K-p} a^\dagger_{1k} a_{1,k+p} \end{aligned} \tag{28}$$

It is then straightforward to use Equation (26) to evaluate the commutator of $\rho_{1p}$ and $\rho_{1,-p'}$ for $p \ge p' > 0$ and obtain

$$[\rho_p, \rho_{-p'}] = \sum_{k=K-p}^{K-p+p'} a^\dagger_{1,k+p-p'} a_{1k} - \sum_{k=-K}^{-K+p'} a^\dagger_{1,k+p-p'} a_{1,k} \tag{29}$$

The right-hand side of this expression has annihilation operators standing to the right of creation operators so it gives zero when acting upon the vacuum state or, when $K \to \infty$, upon any state with a finite number of particles at large momenta. In this sense, the right-hand side of Equation (29) becomes zero when $K \to \infty$. On the other hand, it may be expressed in terms of $\bar{a}_k$, using Equations (26) and (27) to obtain

$$[\rho_{1p}, \rho_{1,-p'}] = \sum_{k=K-p}^{K-p+p'} \bar{a}^\dagger_{1,k+p-p'} \bar{a}_{1k} + \sum_{k=-K}^{-K+p'} \bar{a}^\dagger_{1,k} \bar{a}_{1,k+p-p'} - \delta_{pp'} \sum_{k=-K}^{-K+p'} (1) \tag{30}$$



Once again, as $K \to \infty$, the operator expressions on the right-hand side of this equation give zero in the same sense as before, but the final $c$-number summation tends to a finite limit $-\delta_{pp'} pL/2\pi$ since each $k$ is of the form $2\pi\nu/L$, with $\nu$ an integer. Then as $K \to \infty$ the commutation relation becomes

$$[\rho_{1p}, \rho_{1-p'}] = -\frac{pL}{2\pi} \delta_{pp'} \tag{31}$$

A similar calculation shows that the same result is true for $p' \ge p > 0$ and also that $[\rho_{1p}, \rho_{1p'}] = 0$, $[\rho_{2p}, \rho_{2p'}] = 0$, whereas

$$[\rho_{2p}, \rho_{2,-p'}] = \frac{pL}{2\pi} \delta_{pp'} \tag{32}$$

for $p > 0, p' > 0$. The change in sign occurs because, for the left-going branch, it is the states with $p > 0$ that are filled in the new vacuum. Thus it can be seen that the change in vacuum has a drastic effect on the commutation relations and that, apart from the numerical factor $pL/2\pi$, the density operators are bosons. Indeed $\rho_{1p}$ and $\rho_{2p}$ may be combined into one canonical boson operator $b_p$ by defining

$$\begin{aligned} b^\dagger_p &= i\left(\frac{2\pi}{pL}\right)^{1/2} \rho_{1p}, \qquad p>0 \\ &= -i\left(\frac{2\pi}{|p|L}\right)^{1/2} \rho_{2p}, \qquad p<0 \end{aligned} \tag{33}$$

and the combined commutation relations of the $\rho_{ip}$ become

$$[b_p, b^\dagger_{p'}] = \delta_{pp'} \tag{34}$$

The phase factors $\pm i$ in Equation (33) have been included in anticipation of the requirements of Section 4. Physically $b^\dagger_p$ creates a density wave of momentum $p$ in the original Fermi gas and it is a boson for all $p$, in this model.

The interaction part of $H$ may be expressed immediately in terms of the $b_p$ by substituting Equation (33) into the Fourier transform of Equation (24) to obtain, apart from constants,

$$H_2 = \sum_{p>0} \frac{p}{\pi} [V_1 s (b^\dagger_p b_p + b^\dagger_{-p} b_{-p}) - V_2 s (b^\dagger_p b^\dagger_{-p} + b_{-p} b_p)] \tag{35}$$

In this equation, the operators $b^\dagger_{\pm p} b_{\pm p}$ come from $\rho_{ip}\rho_{i,-p}$ whereas $b^\dagger_p b^\dagger_{-p}$ and its Hermitian conjugate arise from products such as $\rho_{2p}\rho_{1,-p}$ involving one left-going and one right-going density operator.

The representation of the kinetic energy requires more discussion. In terms of the variables $a_k$, Equation (23) may be written

$$H_0 = v_F \sum_k k(a_{1k}^\dagger a_{1k} - a_{2k}^\dagger a_{2k}) \tag{36}$$

and, using Equations (26), (28), and (36), it is straightforward to verify that

$$\begin{aligned}[H_0, \rho_{1p}] &= v_F p \rho_{1p} \\ [H_0, \rho_{2p}] &= -v_F p \rho_{2p}\end{aligned} \tag{37}$$

It is important to notice that these equations, as well as Equation (34), are true only if $k$ ranges from $-\infty$ to $+\infty$ for, if there were a cutoff $K_0$ in $H_0$, all momenta $k > K_0$ would disappear from the left-hand side of Equations (37), and the equations could not be correct since it is necessary to have $K \to \infty$ in Equation (28). It may be that the results obtained by the boson-commutation relations are a good approximation for a system with a large but finite cutoff, but this should be verified in each particular case. Generally speaking, if a cutoff $K$ is included, any physical quantity that is independent of $K$ should be obtained correctly by the boson approach but, overall, the limit $K \to \infty$ must be taken.

The significance of Equations (37) is that, if $|0\rangle$ is the ground state for noninteracting fermions, then $\rho_{ip}|0\rangle$ is an eigenstate with excitation energy $\pm v_F p$, as can be seen by applying both sides of Equation (37) to $|0\rangle$. This is a consequence of the linear dispersion relation used in Equation (36) for, otherwise, it would be true only if $p \to 0$. Similarly it can be seen that other eigenstates may be obtained by acting on $|0\rangle$ with products of the $\rho_{ip}$. Once a set of eigenstates and eigenvalues are known, it is always possible to rewrite the Hamiltonian in terms of them. This may be accomplished here by using Equation (33) to rewrite Equations (37) as

$$[H_0, b_p] = v_F |p| b_p \tag{38}$$

and then, from Equation (34) it is clear that the same equation is satisfied by the *boson representation* of $H_0$:

$$H_{0B} \equiv v_F \sum_p |p| b_p^\dagger b_p + \Delta E \tag{39}$$

where $\Delta E$ is a $c$ number to be determined in a moment. Thus states with any number of charge-density waves are identical to states with the same number of bosons created by the $b_p^\dagger$ and their excitation energies are given exactly by $H_{0B}$. It would be an enormous simplification if $H_{0B}$ could be used in place of $H_0$ for calculating the dynamical and thermodynamic properties of the system, for both $H_{0B}$ and $H_2$ are quadratic forms and hence their sum can be diagonalized by canonical transformation. So far,



however, it is not clear that the boson states are complete, because excited states of the ideal Fermi gas are composed of all allowed combinations and numbers of particle–hole pairs obtained by applying operators such as $a_{p+k}^\dagger a_k$ to $|0\rangle$, and it has not been shown that they are all included in the set of boson states. This question will be addressed in the next section by showing that $\psi_i^\dagger(x)$ can be expressed in terms of the $b_p^\dagger$ and $b_p$, in such a way as to obtain the correct correlation functions.

Some feeling for the validity of using the boson representation may be obtained by evaluating the constant $\Delta E$ appearing in Equation (39). It should merely be the difference of the ground-state energies and should have no thermodynamic significance. This may be checked by equating the free energies $F_0$ and $F_{0B}$ of the ideal Fermi gas and its boson representation. The boson free energy is just the usual expression

$$F_{0B} = \Delta E + k_B T \sum_p \ln(1 + e^{-\beta v_F |p|}) \tag{40}$$

where $T$ is the temperature, $k_B$ is Boltzmann's constant and $\beta = (k_B T)^{-1}$. Changing sums to integrals via $\Sigma_p \to (L/2\pi)\int dp$, the integral may be evaluated to give

$$F_{0B} = \Delta E - \frac{\pi L}{6 v_F}(k_B T)^2 \tag{41}$$

The fermion form of free energy may be evaluated in a similar way, using the operators $\bar{a}_k$ defined in Equation (27), and remembering that $F_{0B}$ refers to a fixed number of particles, since the density operators do not change the number of fermions. The result is

$$F_0 = \frac{L v_F k_F^2}{2\pi} - \frac{\pi L}{6 v_F}(k_B T)^2 \tag{42}$$

Equating $F_{0B}$ and $F_0$ gives $\Delta E = (L v_F k_F^2)/2\pi$, which is just the ground-state energy of the free fermions, and is independent of temperature, as required.

Assuming that $H_B = H_{0B} + H_2$ does represent the interacting system, the thermodynamic properties may be obtained by diagonalization, using a canonical transformation. From Equations (35) and (39), $H_B$ may be written

$$H_B = \sum_{p>0} H_B(p) + \text{const} \tag{43}$$

where

$$H_B(p) = p\left(v_F + \frac{sV_1}{\pi}\right)(b_p^\dagger b_p + b_{-p}^\dagger b_{-p}) - \frac{psV_2}{\pi}(b_p^\dagger b_{-p}^\dagger + b_{-p} b_p) \tag{44}$$

and this quadratic form may be diagonalized by the transformation

$$e^{iS}b_p^\dagger e^{-iS} = b_p^\dagger \cosh\phi_p + b_{-p}\sinh\phi_p$$
$$e^{iS}b_{-p}e^{-iS} = b_p^\dagger \sinh\phi_p + b_{-p}\cosh\phi_p \tag{45}$$

with $\phi_p = \phi_{-p}$. Equations (45) are identical to the transformation used by Bogoliubov[11] in the theory of a Bose gas and, by using the commutation relations (34), it may be verified that

$$S = i\sum_{p>0}(b_p^\dagger b_{-p}^\dagger - b_p b_{-p})\phi_p \tag{46}$$

Using Equations (44) and (45), it can be seen that the off-diagonal terms of the transformed Hamiltonian (those proportional to $b_p^\dagger b_{-p}^\dagger$ and $b_{-p}b_p$) vanish provided

$$\tanh 2\phi_p = \frac{\bar{V}_2}{1+\bar{V}_1} \equiv g \tag{47}$$

with $\bar{V}_i = sV_i/\pi v_F$. Then the full transformed Hamiltonian becomes

$$H_B' \equiv e^{iS}H_B e^{-iS} = \sum_p \varepsilon_p b_p^\dagger b_p + C \tag{48}$$

with

$$\varepsilon_p = v_F|p|[(1+\bar{V}_1)^2 - \bar{V}_2^2]^{1/2} \tag{49}$$

and

$$C = \sum_{p>0}[\varepsilon_p - (1+\bar{V}_1)pv_F] + \Delta E \tag{50}$$

In this equation, the sum over $p$ must be cut off for $p$ larger than $O(s^{-1})$ to obtain a finite result. This reflects the lattice origin of the theory and $C$ is $O(s^{-1})$.

Thus the main effect of the interactions is to modify the energy of the density waves, and the free energy may be obtained from Equation (40) with $v_F|p|$ replaced by $\varepsilon_p$.

As frequently happens in these circumstances, there is a much more significant change in the correlation functions but, in order to understand how this comes about, it is necessary to introduce the boson representations of the $\psi_i(x)$.



### 2.3. Boson Representations of Fermion Fields

The method of evaluating the Luttinger model partition function described in Section 2.2 is essentially the one used by Mattis and Lieb,[10] but the justification is somewhat different and requires the introduction of boson representations of the fermion fields $\psi_i(x)$. The latter also are important in themselves, since they allow us to evaluate correlation functions in a rather direct way and to find special solutions of the model with spin which will be introduced later. Boson representations of $\psi_i(x)$ were introduced first by Schotte and Schotte[12] in order to calculate x-ray threshold transition rates. Later Schotte[13] and also Blume *et al.*[14] applied them to Kondo's model of a magnetic impurity in a metal and derived an expression for the partition function. These problems required a representation of $\psi_i(x)$ at a single point, say $x = 0$, but it is a simple matter to allow for arbitrary values of $x$, either directly or by a gauge transformation, as required for the electron gas in one dimension. Following this route, the correlation functions of Luttinger's model were evaluated by Luther and Peschel,[15] and a solution of the model with spin was obtained for a particular value of a coupling constant by Luther and Emery,[16] who also pointed out the close analogy to the Kondo problem. In describing these developments, the boson representations of $\psi_i(x)$ will be used in a manner that requires a more extensive justification than the original one given by Schotte and Schotte.[12] For this reason, a rather detailed discussion will be given and, in particular, the roles of cutoffs and limiting procedures will be emphasized in order to avoid some of the confusion that has arisen in the literature.

The form of $\psi_i(x)$ is determined almost completely by requiring it to have the correct commutation relations with the density operators. From Equations (25), (26), and (28), it follows that

$$[\psi_i(x), \rho_{ip}] = e^{ipx}\psi_i(x) \tag{51}$$

and, using Equation (33), this equation may be rewritten in terms of boson operators $b_p^\dagger$:

$$[\psi_i(x), b_p^\dagger] = i\frac{2\pi}{|p|L}e^{ipx}\psi_i(x) \tag{52}$$

where $p > 0$ for $\psi_1(x)$ and $p < 0$ for $\psi_2(x)$. These equations are satisfied by the boson form

$$\psi_i = A_i e^{i\Phi_i(x)} \tag{53}$$

where

$$\Phi_1(x) = \sum_{k>0}\left(\frac{2\pi}{|k|L}\right)^{1/2}[b_k^\dagger e^{-ikx} + b_k e^{ikx}]e^{-\alpha|k|/2} \tag{54}$$

and $\Phi_2(x)$ is given by a similar equation with the sum taken over $k<0$. The factor $\exp(-\alpha|k|/2)$ has been included[15] in Equation (55) as a cutoff for divergent integrals which will occur in intermediate steps of the calculations. Physically, $\alpha$ is something like the lattice spacing and there will be numerical factors $\alpha^{1/2}$ replacing $s^{1/2}$ which occurs in Equation (14). Results that are derived from Equations (53) and (54) are only correct if $\alpha \to 0$, and this limit will always be taken. The coefficients $A_i$ in Equation (53) are not determined by Equation (52) since it is homogeneous; but they may be obtained from the fermion commutation relations amongst the $\psi_i(x)$. From Equations (53) and (54),

$$\psi_1^\dagger(x)\psi_1(x') = |A_1|^2 e^{-F^\dagger} e^F \exp\left[\frac{2\pi}{L}\sum_{k>0}\frac{e^{-\alpha k}}{k}(e^{ik(x-x')}-1)\right] \tag{55}$$

where

$$F = \left(\frac{2\pi}{L}\right)^{1/2}\sum_{k>0}\frac{e^{-\alpha k/2}}{k^{1/2}} b_k (e^{ikx}-e^{ikx'}) \tag{56}$$

On the right-hand side of Equation (55), the operators have been normally ordered, i.e., annihilation operators collected into $e^F$ have been moved to the right of creation operators which appear in $e^{-F^\dagger}$. This has been accomplished by using the formulas derived in Appendix A. Changing sums to integrals, the last factor on the right of Equation (55) gives $\alpha/[\alpha - i(x-x')]$ and, evaluating $\psi_1(x')\psi_1^\dagger(x)$ in the same way, it can be seen that

$$\psi_1^\dagger(x)\psi_1(x') + \psi_1(x')\psi_1^\dagger(x) = \alpha|A_1|^2 e^{-F^\dagger} e^F\left[\frac{1}{\alpha - i(x-x')} + \frac{1}{\alpha + i(x-x')}\right] \tag{57}$$

As $\alpha \to 0$, the term in square brackets becomes $2\pi\delta(x-x')$, so it is possible to set $x = x'$ in $F$ and $F^\dagger$ and they vanish. Thus the right-hand side of Equation (57) is a $c$ number, and it is equal to $\delta(x-x')$ provided

$$2\pi\alpha|A_1|^2 = 1 \tag{58}$$

In a similar way it can be shown that $[\psi_1(x), \psi_1(x')]_+ = 0$ and that $\psi_2(x)$ and $\psi_2^\dagger(x')$ satisfy fermion anticommutation relations. Finally it is necessary to arrange for $\psi_1(x)$ and $\psi_1^\dagger(x)$ to anticommute with $\psi_2(x')$ and $\psi_2^\dagger(x')$. This may be accomplished by choosing

$$A_1 = \frac{e^{i\lambda_1 \hat{N}_2}}{(2\pi\alpha)^{1/2}}$$
$$A_2 = \frac{e^{i\lambda_2 \hat{N}_1}}{(2\pi\alpha)^{1/2}} \tag{59}$$



with

$$\hat{N}_i = \int_0^L dx\, \psi_i^\dagger(x)\psi_i(x) \tag{60}$$

and $(\lambda_1 - \lambda_2)$ an odd multiple of $\pi$, as may be verified by direct substitution. Thus, despite the fact that the $\psi_i(x)$ are made up of boson operators, they satisfy fermion anticommutation relations.

As an illustration of the use of this representation of the $\psi_i(x)$, it is instructive to evaluate a single-particle correlation function for the noninteracting system:

$$C_{i,0}(x, t; x', t') = \langle \psi_i(x, t)\psi_i^\dagger(x', t')\rangle \tag{61}$$

where $\langle\cdots\rangle$ denotes thermal average for free fermions or bosons as required, and $\psi_i(x, t)$ is the time evolved operator $e^{iHt}\psi_i(x)\,e^{-iHt}$. Since $C_{i,0}$ is a function of $(x-x')$ and $(t-t')$, it is possible to set $x' = 0$, $t' = 0$ and then, from the fermion form of $\psi_i(x, t)$,

$$C_{1,0}(x, t) = L^{-1}\sum_{k>0} e^{ik(x-v_F t)}(1-f_k) + L^{-1}\sum_{k<0} e^{ik(x-v_F t)} f_k \tag{62}$$

where $f_k$ is the fermion occupation number, $[\exp(\beta v_F|k|)+1]^{-1}$. The right-hand side of Equation (62) may be evaluated by changing sums to integrals to obtain

$$C_{1,0}(x, t) = \frac{1}{2\pi}K(x - v_F t, 0)$$
$$C_{2,0}(x, t) = \frac{1}{2\pi}K^*(x + v_F t, 0) \tag{63}$$

where

$$K(\xi, \eta) \equiv \frac{i}{\xi + i\eta}\frac{\pi\xi/\beta v_F}{\sinh(\pi\xi/\beta v_F)} \tag{64}$$

On the other hand, in the boson picture, Equations (53), (54), (59), and (61) give

$$C_{i,0}(x, t) = \frac{1}{2\pi\alpha}\langle e^{i\Phi_i(x,t)}\, e^{-i\Phi_i(0)}\rangle \tag{65}$$

This is the thermal average of exponentials of linear combinations of creation and annihilation operators and this is the general form that correlation functions will take. It is the principal advantage of the boson representation that such expressions may be evaluated quite simply as

described in Appendix A, and $C_{i,0}(x, t)$ is given by

$$C_{i,0}(x,t)=\frac{1}{2\pi\alpha}\exp-\frac{1}{2}\{\langle[\Phi_i(x,t)-\Phi_i(0,0)]^2\rangle-[\Phi_i(x,t),\Phi_i(0)]\} \tag{66}$$

$$=\frac{1}{2\pi\alpha}\exp\left(\frac{2\pi}{L}\sum_{k>0}\frac{e^{-\alpha|k|}}{k}\{i\ \sin k(x-v_Ft) +(2n_k+1)[\cos k(x-v_Ft)-1]\}\right) \tag{67}$$

where $n_k$ is the boson occupation number. As $L\to\infty$, the sums in this equation become integrals which may be evaluated in closed form to give Equation (63) with $K(\xi,\alpha)$ replaced by

$$\bar{K}(\xi,\alpha)\equiv\frac{i}{\xi+i\alpha}\frac{|\Gamma(1-i\xi/\beta v_F+\alpha/\beta v_F)|^2}{\Gamma^2(1+\alpha/\beta v_F)} \tag{68}$$

The factor $i$ comes from the commutator in Equation (66) and the rest from the thermal average. Letting $\alpha\to 0$ and using $|\Gamma(iy)|^2=\pi/(y\sinh\pi y)$ gives $\bar{K}(\xi,\alpha)=K(\xi,0)$ and the fermion result, Equation (64), is recovered exactly. When $\langle\psi_i(x,t)\psi_i(0,0)\rangle$ is evaluated in the same way, the signs of $\sin k\xi$ and $\cos k\xi$ in Equation (67) are changed and the integral diverges at the lower limit to give zero for the correlation function, as it should.

This method can easily be extended to evaluate expectation values of products of any number of $\psi_i(x,t)$ and $\psi_i^\dagger(x',t')$. Using Equations (53) and (54) and the results of Appendix A, the various factors may be combined into one exponential of the form $\exp[i\Sigma_n(\pm)\Phi_i(x_n,t_n)]$ and, as in Equation (66), the correlation function becomes the exponential of a sum of pair functions. This is precisely the product of pair functions which is obtained directly from the initial fermion variables, and the correct overall sign is guaranteed because the anticommutation relations are satisfied.

These results are sufficient to demonstrate that the boson representation is correct for the interacting system, at least in perturbation theory, since, order by order, the correlation function or the partition function are integrals of a product of free-particle correlation functions. This is where the true value of the boson representation is realized, for it enables us to obtain results for the interacting system with much greater ease than by using the fermion variables directly.

Finally, another view of Equations (53) and (54) may be obtained by differentiating $\Phi_i(x)$ and using Equations (33) to find

$$\frac{d\Phi_i(x)}{dx}=(-1)^{i+1}2\pi\rho_i(x) \tag{69}$$

so that $\Phi_i$ is the integral of a physical quantity, the density. In other words, $\Phi_i$ is a nonlocal quantity and it is essentially this feature that allows $\psi_i(x)$ to



satisfy anticommutation relations. Also, from Equations (55), (56), (58), and (69),

$$\psi_1^\dagger(x)\psi_1(x')=-\frac{1}{2\pi i}\frac{1}{x-x'+i\alpha}+\rho_1(x) \tag{70}$$

for $x\approx x'$. The first term is undefined as $x\to x'$ and $\alpha\to 0$, but it has to be there in order to give the $\delta$ function in the anticommutation relation (57), and it is the usual infinite term obtained from the product of two quantum fields at the same point. There is no inconsistency, since the singular term contributes to the $p=0$ component of the Fourier transform of $\rho_i(x)$ and this does not occur in Equation (33). For $p\neq 0$, the correct relation between $\rho_{ip}$ and $b_p$ is obtained. Indeed Equations (53) and (54) could have been used as a starting point for the discussion and the physical interpretation of the $b_p$ deduced from that.

### 2.4. *Correlation Functions of the Interacting System*

In this section, the boson representations will be used to evaluate the one-particle Green's function and various two-particle correlation functions from which the physical properties of Luttinger's model may be obtained. The interesting physical questions concern the existence of a Fermi surface and the tendency towards superconducting or charge-density wave order as the temperature is decreased, but the mathematical results will also be of use in considering fermions with spin. The correlation functions have also been evaluated by directly using fermion variables,[17] but it will be seen that the boson representations, which were used for this purpose first by Luther and Peschel,[15] give a very simple method of calculation.

The first quantity to be evaluated will be the retarded one-particle Green's function

$$G_i^R=-i\theta(t-t')\langle[\psi_i(x,t),\psi_i^\dagger(x',t')]_+\rangle \tag{71}$$

where $\theta(t)$ is the unit step function. Translational invariance implies that $G_i^R$ is a function of $(x-x')$ and $(t-t')$ so it is possible to set $x'=0$, $t'=0$ without loss of generality. It can be seen that $G_i^R(x,t)$ contains the correlation function $C_i(x,t)$ defined in Equation (61) and, for a noninteracting system, using Equations (63) and (64) together with Equation (71),

$$G_i^R(x,t)=-2i\theta(t)\,\mathrm{Re}\,C_{i,0}(x,t)$$

$$=\begin{cases}-i\theta(t)\delta(x-v_Ft), & i=1\\ -i\theta(t)\delta(x+v_Ft), & i=2\end{cases} \tag{72}$$

In the presence of interactions it is necessary to diagonalize the Hamiltonian with the aid of the transformation (45) and, to obtain $G_i^R(x,t)$, it is merely necessary to find the effect on $\Phi_i(x)$. From Equations (45) and (54),

$$\Phi_1'(x)=e^{iS}\Phi_1(x)\,e^{-iS}=\sum_{k>0}\left(\frac{2\pi}{|k|L}\right)^{1/2}\cosh\phi_k\,[b_k^\dagger e^{-ikx}+b_k e^{ikx}]\,e^{-\alpha|k|/2}$$
$$+\sum_{k<0}\left(\frac{2\pi}{|k|L}\right)^{1/2}\sinh\phi_k\,[b_k^\dagger e^{-ikx}+b_k e^{ikx}]\,e^{\alpha|k|/2} \tag{73}$$

Similarly $\Phi_2'(x)$ is obtained by interchanging $\sinh\phi_k$ and $\cosh\phi_k$. For practical calculations, it is simplest to assume that $V_p$ and $\phi_p$ are constants $V$ and $\phi$ and then from Equations (47) and (49), $\varepsilon_p=v_F'|p|$ and $\sinh^2\phi\equiv\gamma$, where

$$v_F'=v_F[(1+\bar V_1)^2-\bar V_2^2]^{1/2} \tag{74}$$

and

$$2\gamma=(1-g^2)^{-1/2}-1 \tag{75}$$

However, to obtain finite values of integrals, it is necessary to introduce a cutoff. This does not spoil the boson representation and it is the one allowed way of introducing a cutoff in using this method. Conceptually, the simplest assumption is to take $V_p=V$ for $|p|\le p_0$ and $V=0$ for $|p|>p_0$, but this can introduce awkward nonanalyticities, and the alternative form $\sinh^2\phi_p=\gamma e^{-rp}$, used by Luther and Peschel,[15] is much more convenient and will be adopted here. Wherever possible, $r$ will be set equal to zero and then it need not be introduced into Equation (74), serving mainly to define otherwise singular quantities.

Using Equation (73) and the corresponding one for $\Phi_2(x)$, it can be seen that, when $\phi_k$ is a constant, the canonical transformation actually mixes up the correlation functions $C_{1,0}(x,t)$ and $C_{2,0}(x,t)$. From this fact, the Green's function for the interacting system can be written down immediately in terms of that of the noninteracting system, as in Equations (63), (66)–(68). From Equation (73) the contribution from $k>0$ to Equation (66) will involve $\cosh^2\phi_p=1+\sinh^2\phi_p=1+\gamma e^{-rp}$. The first term (for $\gamma=0$) will lead to $C_{1,0}(x,t)$ and the second term will give $C_{1,0}^\gamma(x,t)$ with $\alpha$ replaced by $r$, since $e^{-rp}$ takes over the role of cutoff and $\alpha$ may be set equal to zero. Similarly the $k<0$ contribution from Equation (73) gives $C_{2,0}^\gamma(x,t)$. In all of these functions, $v_F$ is replaced by $v_F'$. The net result is that

$$G_1^R(x,t)=-\frac{i\theta(t)}{\pi}\,\mathrm{Re}\{K(x-v_F't,0)[r^2K(x-v_F't,r)K^*(x+v_F't,r)]^\gamma\} \tag{76}$$



where $K(\xi,\alpha)$ is given in Equation (64). To obtain $G_2^R(x,t)$, it is necessary to interchange $x+v_F't$ and $x-v_F't$. It can be seen that, for large $x$, $G_1^R(x,t)$ falls off exponentially with a correlation length proportional to $T^{-1}$, whereas, at $T=0$, it becomes a power law:

$$G_1^R(x,t)\underset{T\to0}{\to}i\theta(t)\,\mathrm{Im}\left\{\frac{1}{x-v_F't+i0}\left[\frac{r^2}{(x-v_F't+ir)(x+v_F't-ir)}\right]^\gamma\right\} \tag{77}$$

This result is exact for Luttinger's model but it is also correct asymptotically for a lattice model and is independent of the details of the cutoff. It is characteristic of the boson method that, in contrast to the usual methods of many-body theory, the correlation functions are obtained most naturally in real space and time. The Fourier transformation of $G_i^R(x,t)$ for $T=0$ is described by Luther and Peschel,[15] and, in particular, they find that

$$-\frac{1}{\pi}\,\mathrm{Im}\,G_1^R(q=0,\omega+i0)\simeq\frac{r}{v_F'}\left(\frac{\omega r}{2v_F'}\right)^{\gamma-1}\Gamma(1-\gamma)\sin\pi\gamma/2$$
$$+\frac{2r}{\gamma v_F'}\left(\frac{\omega r}{2v_F'}\right)^{2\gamma-1}[\Gamma(1-\gamma)\sin\pi\gamma]^2 \tag{78}$$

for $\omega\ll v_F'/r$. This is to be compared to the usual form $Z\delta(\omega)+b(\omega)$ obtained in three dimensions, where the $\delta$ function comes from the quasiparticle pole. It is clearly absent from Equation (78) and, as first pointed out by Gutfreund and Schick,[18] there are no fermion quasiparticles in Luttinger's model. In this sense, the fermions are "confined," and this is true for all $\omega$ when $r\to0$. This is consistent with the conclusion that the Hamiltonian may be expressed completely in terms of density waves. If the right-hand side of Equation (77) is expanded in powers of $\gamma$, every order will contain powers of $\ln\omega$, which diverges as $\omega\to0$, and they must be summed to obtain an accurate result. This feature is common to all one-dimensional systems and is not a consequence of the continuum limit. It is clear that selective summations, such as the random-phase approximation, do not give correct results in these circumstances.

The other interesting correlation functions are the $2k_F$ susceptibility

$$\chi^R(x,t)=-i\theta(t)\langle[\psi_1^\dagger(x,t)\psi_2(x,t),\psi_2^\dagger\psi_1]\rangle \tag{79}$$

describing the response of the system to a perturbation which transfers an electron from the neighborhood of one fermi point to the other (i.e., right-going fermions to left-going fermions and vice versa) and the pairing correlation function

$$P^R(x,t)=-i\theta(t)\langle[\psi_1(x,t)\psi_2(x,t),\psi_2^\dagger\psi_1^\dagger]\rangle \tag{80}$$

Physically, divergences in the Fourier transforms $\chi^R(\omega, 0)$ or $P^R(\omega, 0)$ are a sign that a many-chain system may undergo a transition to a CDW state or a superconductive state as described in Section 1.

The procedure for evaluating $\chi^R(x, t)$ and $P^R(x, t)$ is exactly the same for the single-particle correlation functions. Introducing the boson representation for $\psi_i(x)$ gives an expression similar to that in Equation (66) except that a product of four exponentials has to be evaluated and the mixing of $\Phi_1(x)$ and $\Phi_2(x)$ leads to factors $\cosh\phi_p \pm \sinh\phi_p = \exp(\pm\phi_p)$ which replace the separate $\cosh\phi_k$ and $\sinh\phi_k$ factors in Equation (73). Then it is convenient to introduce a slightly different cutoff $\exp(\pm 2\phi_p) \to 1 + [\exp(\pm 2\phi_p) - 1]\exp(-rp)$, which again enables integrals to be evaluated simply. As a result

$$\chi^R(x, t) = \frac{\theta(t)}{2\pi^2 r} \operatorname{Im}[r^2 K(x - v_F' t, r) K^*(x + v_F' t, r)]^{\mu} \tag{81}$$

and

$$P^R(x, t) = \frac{\theta(t)}{2\pi^2 r} \operatorname{Im}[r^2 K(x - v_F' t, r) K^*(x + v_F' t, r)]^{1/\mu} \tag{82}$$

where

$$\mu = e^{-2\phi} = \left(\frac{1-g}{1+g}\right)^{1/2} \tag{83}$$

and $g$ is given by Equation (47). Using Equation (83) for $\mu$, it can be seen that $\chi^R(x, t)$ and $P^R(x, t)$ differ only by the sign of $g$. The Fourier transforms of $\chi^R$ and $P^R$ have been evaluated by Luther and Peschel.[15] For $T = 0$, they find

$$\operatorname{Im}\chi^R(q, \omega) = \begin{cases} 0, & |\omega| < v_F'|q| \\ (\pi^2/v_F')\Gamma^2(1-\gamma)\sin^2\pi(1-\gamma) & \\ \quad \times [(r^2/4v_F'^2)(\omega^2 - c^2 q^2)]^{\mu-1}, & |\omega| > v_F'|q| \end{cases} \tag{84}$$

Here, because of the phase factors which have been removed in Equations (21) and (22), the wave vector $q$ is actually measured relative to $2k_F$. Again there is a power law singularity at the edge of the continuum, which can only be found by summing all orders of perturbation theory. There is a divergence for $\mu < 1$ or $g > 0$ (corresponding to repulsive interactions) and this will lead to a CDW instability for coupled chains at sufficiently low temperatures. On the other hand, for attractive interactions, $\mu > 1$ and $\chi(q, \omega)$ vanishes as $\omega \to 0$, $q \to 0$. The pairing correlation function $P^R(q, \omega)$ is obtained by changing $\mu$ to $\mu^{-1}$ in Equation (83) and this is divergent



when $g < 0$. Here the phase factors of Equation (20) cancel and $q$ refers to the total momentum of a pair. Physically, a divergence in $P(q, \omega)$ implies a tendency towards superconductive pairing when there is an appropriate coupling between chains.

These are the qualitative features of the correlation function of Luttinger's model. A more extensive discussion may be found in the literature.[15,17] However, to get closer to the real physical situation, it is necessary to take account of the electron spin, and this produces a qualitative change in the physics and in the mathematics that are required to describe it. The problem becomes much more difficult because the Hamiltonian, now given by Equations (12)–(15), can no longer be written as a quadratic form in boson variables, and the remaining sections are concerned with the various limiting situations in which some progress can be made. They have the common feature that the charge and spin degrees of freedom separate, but in a way that usually does not allow an explicit solution to be found as simply as for Luttinger's model.

## 3. Large "On-Site" Interaction

The simplest situation for fermions with spin occurs when the "on-site" interaction $U$, defined in Equation (5), is the largest energy in the problem, and it will be considered in some detail, for not only is it the basis of the simple physical picture described in Section 1, but it may be relevant for real materials in which there are strong Coulomb forces.[19] To be complete, it is necessary to consider $U < 0$ also, and this has some pedagogical value for, even if large negative $U$ is unlikely to be found in nature, a feeling for what happens at smaller negative values is useful, because they may well arise when electrons delocalize on large molecules (thereby reducing the Coulomb force) and attractive couplings are induced by molecular polarizability or other collective effects.[7]

The method of calculation will be to regard $H_0$ and $H_2$ of Equations (4) and (6) as perturbations, and to work to the lowest nontrivial order. As described in Section 1, the unperturbed states depend upon the sign of $U$.

### 3.1. Attractive Interaction

Here the unperturbed ground state has $N_e/2$ of the sites occupied by two opposite-spin electrons. These are tightly bound pairs which replace the Cooper pairs[4] occurring for weak coupling, and they are responsible for superconductivity. The excited states have one or more pairs broken and are separated in energy by multiples of $|U|$. When $N_e/2 < M$, some

sites are unoccupied and there is a substantial degeneracy because the energy does not depend upon how they are distributed, but the inclusion of $H_0$ and $H_2$ splits the levels into nonoverlapping bands of charge-density waves in which pairs of electrons move from site to site. Interband transitions with one or more pairs broken correspond to spin-density wave excitations. The discussion will be restricted to the lowest band, which is sufficient for a calculation of the properties of the system at low temperatures.

In first-order degenerate perturbation theory, $H_2$ has to be diagonalized in the space of paired states and, within that space, it is an effective Hamiltonian from which the thermodynamic properties may be obtained. But $H_0$ breaks pairs and so has no first-order matrix elements. It must therefore be calculated to second order, allowing virtual transitions into the next band of states which have one pair broken. Suppose $H_2 = 0$ for the moment, and let the various degenerate ground states of $H_1$ be denoted by $|\alpha\rangle$ with energy $E_1 = NU/2$. The Schrödinger equation is

$$(E - H_1)|\Psi\rangle = H_0|\Psi\rangle \tag{85}$$

and dividing both sides by $E - H_1$ and rearranging gives

$$|\Psi\rangle = \sum_\alpha |\alpha\rangle \frac{\langle\alpha|H_0|\Psi\rangle}{E - E_1} + \frac{P}{E - H_1} H_0|\Psi\rangle \tag{86}$$

where $P = 1 - |\alpha\rangle\langle\alpha|$ projects out of the unperturbed ground states. By substitution, it can be seen that Equation (85) is equivalent to

$$|\Psi\rangle = \sum_\alpha a_\alpha |\Psi_\alpha\rangle \tag{87}$$

where

$$|\Psi_\alpha\rangle = |\alpha\rangle + [P/(E - H_1)]H_0|\Psi_\alpha\rangle \tag{88}$$

and

$$a_\alpha = \langle\alpha|H_0|\Psi\rangle/(E - E_1) \tag{89}$$

To first order, Equation (88) is

$$\begin{aligned} |\Psi_\alpha\rangle &= |\alpha\rangle + [P/(E - H_1)]H_0|\alpha\rangle \\ &= |\alpha\rangle(1/U)H_0|\alpha\rangle \end{aligned} \tag{90}$$

The last line follows because $H_0$ breaks exactly one pair to give an excitation energy $-U$ and $P$ is irrelevant because $H_0|\alpha\rangle$ has no component in the ground states. Then, substituting Equations (87) and (90) into Equation (89) gives

$$(E - E_1)a_\alpha = \frac{1}{U}\sum \langle\alpha|H_0^2|\alpha'\rangle a_{\alpha'} \tag{91}$$



which is a Schrödinger equation in the $|\alpha\rangle$ subspace with effective Hamiltonian $H_0^2/U$. To ensure that $H_0$ acts between the ground states, it is necessary that, if the first application of $H_0$ transfers an electron with spin $\sigma$ from site $n$ to site $m$, then the second application of $H_0$ either returns the electron to its original site or transfers the electron with spin $(-\sigma)$ from $n$ to $m$. Thus the effective Hamiltonian is

$$H_0' = -\frac{\varepsilon^2}{|U|}\sum_{n,\sigma} (\psi^\dagger_{n\sigma}\psi_{n+1,\sigma}\psi^\dagger_{n+1,\sigma}\psi_{n\sigma} + \psi^\dagger_{n,-\sigma}\psi_{n+1,-\sigma}\psi^\dagger_{n\sigma}\psi_{n+1,\sigma}) + \text{H.c.} \tag{92}$$

Since it is sufficient to work to first order in $H_2$, the total effective Hamiltonian is

$$H' = H_0' + H_2 \tag{93}$$

which may be used for temperatures $T \ll |U|$ where the probability of having a broken pair is very small.

A feeling for the physics represented by this Hamiltonian may be obtained by introducing boson variables that are quite different from those used in Section 2. Define

$$\rho_n = \tfrac{1}{2}(\rho_{n\uparrow} + \rho_{n\downarrow} - 1) \tag{94}$$

$$b_n = \psi_{n\uparrow}\psi_{n\downarrow} \tag{95}$$

$$\sigma_n = \tfrac{1}{2}(\rho_{n\uparrow} - \rho_{n\downarrow}) \tag{96}$$

where $\rho_{n\sigma} = \psi^\dagger_{n\sigma}\psi_{n\sigma}$ as in Equation (6). Then

$$\begin{aligned} H' = \frac{2\varepsilon^2}{|U|}\sum_n (-b_n^\dagger b_{n+1} - b_{n+1}^\dagger b_n + 2\rho_n\rho_{n+1} + 2\sigma_n\sigma_{n+1}) \\ + V\sum_m (2\rho_m + 1)(2\rho_{m+1} + 1) + \text{const} \end{aligned} \tag{97}$$

The operators $\sigma_n$ commute with the $b_n$ and $\rho_n$ and give zero when applied to an empty site or a doubly occupied site, so they may be ignored for the lowest band. The variables $b_m$ and $\rho_n$ satisfy the commutation relation

$$\begin{aligned} [b_m, b_n^\dagger] = [b_m, b_n] = 0, \qquad m \neq n \\ b_m^2 = 0 \\ [\rho_m, b_n^\dagger] = b_n^\dagger \delta_{mn} \end{aligned} \tag{98}$$

which imply that they may be interpreted as boson operators and their number operators, provided there is a hard-core interaction preventing two or more particles from occupying a single site. The question of the existence of superconductivity in the original system is now rephrased as

the existence of a condensation of the bosons and, as mentioned in Section 1, this is quite independent of the formation of pairs, a distinction that persists for weak coupling, in contrast to the behavior of more isotropic materials, where the BCS picture[4] takes over as the coupling is reduced. In fact, when the correlation functions are evaluated, it will become clear that quantum fluctuations prevent Bose condensation even in the ground state of a one-dimensional system and the existence of an energy gap is a rather local phenomenon related to bound pairs rather than to long-range order.

### 3.2. Repulsive Interaction

Here the hopping term $H_0$ breaks the degeneracy in first order, and although it is quite easy to deal with the translational motion, it is much more difficult to incorporate the spin degrees of freedom. An exception is the half-filled band, $N_e = M$ for which electron hopping leads to doubly occupied sites. The effective Hamiltonian for this case can be obtained from Equation (97) by using the canonical transformation

$$\begin{aligned} \tilde{\psi}^{\dagger}_{m\downarrow} &= (-1)^m \psi_{m\downarrow} \\ \tilde{\psi}_{m\uparrow} &= \psi_{m\uparrow} \end{aligned} \tag{99}$$

which does not affect $H_0$, but changes the sign of $U$ in Equation (5), while adding a multiple of the number operator. In the unperturbed state, sites that contained a down-spin electron are now empty, whereas the other sites are doubly occupied. Thus the path traced out for $U<0$ may be followed once again and only the physical significance of the variables is changed—now the boson operators of Equations (94)–(96) represent spin-density waves, and the energy gap immobilizes the charge-density degrees of freedom at low temperatures. This interchange of the roles of charge- and spin-density waves is not peculiar to strong coupling and appears in the continuum limit to be considered in Section 4.

Since $\Sigma_\sigma \rho_{m\sigma}$ is unity for every site $m$, it follows that $H_2$ is a constant in lowest order and the effective Hamiltonian is given by Equation (97) with $V = 0$.

### 3.3. Correlation Functions

The bosons of Equation (94)–(98) are physically appealing, but the hard-core interaction makes it difficult to evaluate the correlation functions. A more useful representation may be obtained by noticing that Equation (98) are the commutation relations of spin-half operators



$S^x_m, S^y_m, S^z_m$ with the identification

$$\begin{aligned} b_m &= S^-_m \\ b^{\dagger}_m &= S^+_m \\ \rho_m &= S^z_m \end{aligned} \tag{100}$$

where $S^{\pm}_m = S^x_m + iS^z_m$. In this notation, the Hamiltonian may be rewritten

$$H' = -\sum_m \left(J_x S^x_m S^x_{m+1} + J_y S^y_m S^y_{m+1} + J_z S^z_m S^z_{m+1}\right) \tag{101}$$

where

$$\begin{aligned} J_x &= J_y \\ &= 4\varepsilon^2/|U| \\ J_z &= -4\varepsilon^2/|U| - 4V \end{aligned} \tag{102}$$

with $V = 0$ for $U > 0$. It is simplest to carry out calculations in the grand canonical ensemble for the original problem, and since the number operator is essentially $\Sigma_m S^z_m$, the chemical potential becomes an effective magnetic field. The discussion here will be restricted to zero field, corresponding to a half-filled band.

As described in Section 2, the trend to various kinds of long-range order is expressed in the correlation functions generated by $\psi^{\dagger}_{m\sigma}\psi_{n\sigma}$ or $\psi_{m\sigma}\psi_{m\sigma}$. When $U<0$, the wave functions for the lowest band are linear combinations of the $|\Psi_\alpha\rangle$ of Equation (90) and they, in turn, are linear combinations of states with all sites occupied by pairs of electrons of opposite spin or by pairs broken without spin-flip. Only $\psi^{\dagger}_{m\uparrow}\psi_{n\uparrow}$, $\psi^{\dagger}_{m\downarrow}\psi_{n\downarrow}$ or $\psi^{\dagger}_{m\uparrow}\psi^{\dagger}_{n\downarrow}$ have matrix elements within this space, and all other combinations connect to states separated from the ground state by an energy gap which prevents divergences at low frequency. It is sufficient to use the bare operators to generate correlation functions, since the admixture of broken pairs in $|\Psi_\alpha\rangle$ is of order $|\varepsilon/U|$, so $|\Psi_\alpha\rangle \approx |\alpha\rangle$ and the role of $H_0$ and $H_2$ is merely to determine the coefficients $a_\alpha$ that appear in Equation (87). From this and Equations (94), (95), and (100), it follows that only the operators

$$\tfrac{1}{2}(\psi^{\dagger}_{m\uparrow}\psi_{m\uparrow} + \psi^{\dagger}_{m\downarrow}\psi_{m\downarrow}) \equiv S^z_m + \tfrac{1}{2} \tag{103}$$

and

$$\psi^{\dagger}_{m\uparrow}\psi^{\dagger}_{m\downarrow} \equiv S^+_m \tag{104}$$

need be considered. In the pseudospin picture, for $U<0$, a charge-density wave state corresponds to antiferromagnetic order in the $z$ direction and superconductivity to ferromagnetic order in the $x$–$y$ plane. The analogous picture for $U>0$ may be obtained by substituting Equations (99) into

Equations (103) and (104), leading to operators that generate longitudinal and transverse spin-density waves, respectively.

For $|J_z/J_x| \leq 1$, the asymptotic forms of the correlation functions have been obtained by Luther and Peschel,[20] who combined results known for the lattice model with scaling laws derived for the continuum limit to find that, if $x = ms$ and $T = 0$,

$$\langle S^z(x,t)S^z\rangle = (2\pi^2 r^2)^{-1}(\cos 2k_F x)\left(\frac{r^2}{x^2 - v_F^2 t^2}\right)^{1/2\theta} \tag{105}$$

$$\langle S^+(x,t)S^-\rangle + \langle S^-(x,t)S^+\rangle = (2\pi^2 r^2)^{-1}\left(\frac{r^2}{x^2 - v_F^2 t^2}\right)^{\theta/2} \tag{106}$$

Here, $r$ is the cutoff introduced in Section 2 and

$$\theta = \tfrac{1}{2} - \pi^{-1}\arcsin(J_z/J_x) \tag{107}$$

For large $x$, the right-hand sides of Equations (105) and (106) are proportional to $x^{-1/\theta}$ and $x^{-\theta}$, respectively, and, since $0 \leq \theta \leq 1$ for the range of $J_z/J_x$ under consideration, it follows that there is no long-range order in the ground state—a consequence of quantum fluctuations rather than the classical thermal fluctuations described in Section 1. If there is an appropriate coupling between the one-dimensional segments of the system, the existence of a phase transition requires much less of the correlation functions. Their Fourier transforms are proportional to $\omega^\mu$ [with $(\mu = \theta^{-1} - 2, q = 2k_F)$ for Equation (105) and $(\mu = \theta - 2, q = 0)$ for Equation (106)], and merely have to be sufficiently large at low frequency,[21] which is true if $\mu < 0$. The resulting physical picture is precisely the one set forth in Section 1.

A more detailed account of the strong coupling limit may be found in the literature.[21–23] In particular, a discussion of the region $J_z < -J_x = -J_y$ [or, by Equation (102) $V > 0, U < 0$] has been given by Fowler,[23] who finds that the ground state has long-range charge-density wave order and that there is a gap in the charge-density spectrum as well as the one in the spin-wave spectrum discussed above. This situation may also be realized for weaker coupling, as will be seen in Section 4.

## 4. Continuum Limit—Energy Gaps

In this section, the continuum limit will be introduced at the outset with a view to obtaining the long-range behavior of the correlation functions in much the same way as for spinless fermions. It turns out, however, that the introduction of spin produces a number of significant modifications. As might be expected from the discussion of the strong coupling limit, the spectra may now have energy gaps, and this is related to the fact that the Hamiltonian is no longer quadratic in boson variables as it was in Section 2. Moreover, the continuum limit does not exist for all values of the coupling constant, since ultraviolet divergences may drive the free energy to minus infinity. Fortunately the use of boson representations is nicely supplemented by the renormalization group method (Section 5), in that each works where the other fails, but it will be necessary to wait until the end of Section 5 for a full discussion of the possible phases of the system.



### 4.1. Separation of Charge and Spin Degrees of Freedom

If the original Hamiltonian of Equations (3)–(6) is rewritten in terms of the right-going and left-going operators $\psi_{1\sigma}$ and $\psi_{2\sigma}$, the boson variables that may then be introduced acquire a spin label, and a significant simplification is achieved by introducing charge- and spin-density operators

$$\begin{aligned} c_p^\dagger &= (b_{p\uparrow}^\dagger + b_{p\downarrow}^\dagger)/2^{1/2} \\ s_p^\dagger &= (b_{p\uparrow}^\dagger - b_{p\downarrow}^\dagger)/2^{1/2} \end{aligned} \tag{108}$$

in terms of which the problem will prove to separate into two independent parts. First of all the quadratic, Luttinger model contribution $H_{LM}$ must have the general form

$$\begin{aligned} H_{LM} = &\sum_{p>0}\left[v_c p(c_p^\dagger c_p + c_{-p}^\dagger c_{-p}) - \frac{sW_\parallel}{2\pi}(c_p^\dagger c_{-p}^\dagger + c_{-p}c_p)\right] \\ &+ \sum_{p>0}\left[v_s p(s_p^\dagger s_p + s_{-p}^\dagger s_{-p}) + \frac{sU_\parallel}{2\pi}(s_p^\dagger s_{-p}^\dagger + s_{-p}s_p)\right] \end{aligned} \tag{109}$$

in which there are no products such as $s_p^\dagger c_{-p}$, involving one charge and one spin variable, because the Hamiltonian should be invariant under reversal of the spins which, by Equation (108) changes $s_p$ to $-s_p$ and leaves $c_p$ unchanged. As a consequence, the charge- and spin-density waves are

*Table 1. Relations between Coupling Constants*

| Quantity | "g-ology" picture | Extended Hubbard model |
|---|---|---|
| $v_c$ | $v_F + g_4/\pi$ | $v_F + sU/2\pi + 2sV/\pi$ |
| $W_\parallel$ | $(2g_2 - g_1)/s$ | $U + 6V$ |
| $W_\perp$ | $g_3/s$ | $U - 2V$ |
| $v_s$ | $v_F$ | $v_F - sU/2\pi$ |
| $U_\parallel$ | $g_1/s$ | $U - 2V$ |
| $U_\perp$ | $g_1/s$ | $U - 2V$ |

uncoupled in $H_{LM}$. Table 1 shows the relationship between the notation of Equation (109) (which reflects the differing roles of various terms in determining the solution) and that of the original Hamiltonian. Also included is the so-called "$g$-ology" notation, which has frequently been used in this context.[25]

The other parts of the Hamiltonian which come from taking the continuum limit are backward scattering

$$H_{BS} = sU_\perp \int dx\, \psi^\dagger_{1\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{2\uparrow} + \text{H.c.} \tag{110}$$

in which electrons of opposite spin cross the Fermi surface in *opposite* directions, and umklapp scattering

$$H_U = sW_\perp \int dx\, e^{4i\bar{k}_F x}\psi^\dagger_{2\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{1\uparrow} + \text{H.c.} \tag{111}$$

in which electrons of opposite spin cross the Fermi surface in the *same* direction. In Equation (111), $4\bar{k}_F \equiv 4k_F - G$ and the momentum transfer is absorbed completely by the lattice for electrons at the Fermi surface when there is a half-filled band $G = 4k_F$. In the absence of spin, $H_U$ vanishes [because $\psi_1^2(x) = 0$] and, by permuting operators, the integrand in $H_{BS}$ may be rewritten as $\rho_1(x)\rho_2(x)$ and incorporated into the Luttinger model Hamiltonian. [In fact this is why $V_2 = 2V$ rather than just $V$ in Equation (24).] Apart from the factor $s$, the notation for the coupling constants is that of Luther and Emery[16] and of Emery *et al.*[24] The relation to other descriptions is given in Table 1.

In order to obtain a boson representation of $H_{BS}$ and $H_U$, it is necessary to add a spin label to Equations (53) and (54); then, using Equation (108) it is found that

$$\psi^\dagger_{1\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{2\uparrow} = (2\pi\alpha)^{-2}\exp[-2^{1/2}i(\Phi_{1s} - \Phi_{2s})] \tag{112}$$

and

$$\psi^\dagger_{2\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{1\uparrow} = (2\pi\alpha)^{-2}\exp[2^{1/2}i(\Phi_{1c} - \Phi_{2c})] \tag{113}$$

where $\Phi_{is}$ and $\Phi_{ic}$ are obtained by substituting $s_p$ and $c_p$, respectively, for $b_p$ in Equation (54). From Equations (112) and (113) it is clear that backward scattering affects only the spin-density waves and that umklapp scattering affects only the charge-density waves, so the Hamiltonian completely separates for the two kinds of collective mode.

### 4.2. Reduction to Spinless Fermions

The possibility of obtaining a solution now relies upon the way in which the right-hand sides of Equations (112) and (113) behave under the canonical transformation (45) and (46)

$$e^{iS}(\Phi_{1\lambda} - \Phi_{2\lambda})\, e^{-iS_0} = e^{-\phi}(\Phi_{1\lambda} - \Phi_{2\lambda}) \tag{114}$$

where $\lambda$ represents either the $s$ or $c$ subscripts. Then by choosing $\phi = (\ln 2)/2$, it is possible to transform the right-hand sides of Equation (112) to $\exp[-i(\Phi_{1s} - \Phi_{2s})]$ and Equation (113) to $\exp[i(\Phi_{1c} - \Phi_{2c})]$, and from Equation (53) they are boson representations of products of *two spinless* fermion operators:

$$e^{iS}\psi^\dagger_{1\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{2\uparrow}\, e^{-iS} = (2\pi\alpha)^{-1}\psi^\dagger_{1s}\psi_{2s} \tag{115}$$

and

$$e^{iS}\psi^\dagger_{2\uparrow}\psi^\dagger_{2\downarrow}\psi_{1\downarrow}\psi_{1\uparrow}\, e^{-iS} = (2\pi\alpha)^{-1}\psi^\dagger_{2c}\psi_{1c} \tag{116}$$

Notice that the phase factors involving fermion number operators, which must be introduced to ensure the full anticommutation relations as in Equation (59), do not appear in Equations (115) and (116), as shown in Appendix B.

If the canonical transformation is now applied to the remainder of the Hamiltonian and the result is rewritten in terms of the spinless fermion variables $\psi_{1c}$ and $\psi_{2c}$, it is found that

$$e^{iS}H\,e^{-iS} = H_s + H_c \tag{117}$$

where

$$\frac{H_s}{v_s} = \frac{i}{8}(10 + 3\bar{U}_\parallel)\int_0^L dx\left(\psi^\dagger_{1s}\frac{\partial\psi_{1s}}{\partial x} - \psi^\dagger_{2s}\frac{\partial\psi_{2s}}{\partial x}\right) - \frac{\pi}{4}(6 + 5\bar{U}_\parallel)\int_0^L dx\,\rho_{1s}\rho_{2s} + \frac{\bar{U}_\perp}{2\alpha}\int_0^L dx\,(\psi^\dagger_{2s}\psi_{1s} + \psi^\dagger_{1s}\psi_{2s}) \tag{118}$$

and

$$\frac{H_c}{v_c} = \frac{i}{8}(10 - 3\bar{W}_\parallel)\int_0^L dx\left(\psi^\dagger_{1c}\frac{\partial\psi_{1c}}{\partial x} - \psi^\dagger_{2c}\frac{\partial\psi_{2c}}{\partial x}\right) - \frac{\pi}{4}(6 - 5\bar{U}_\parallel)\int_0^L dx\,\rho_{1c}\rho_{2c} + \frac{\bar{W}_\perp}{2\alpha}\int_0^L dx\,(e^{4i\bar{k}_F x}\psi^\dagger_{2c}\psi_{1c} + e^{-4i\bar{k}_F x}\psi^\dagger_{1c}\psi_{2c}) \tag{119}$$

with

$$\bar{U}_\parallel = \frac{sU_\parallel}{\pi v_s}, \quad \bar{U}_\perp = \frac{sU_\perp}{\pi v_s}, \quad \bar{W}_\parallel = \frac{sW_\parallel}{\pi v_c}, \quad \bar{W}_\perp = \frac{sW_\perp}{\pi v_s} \tag{120}$$

As shown in Section 2, the bare Fermi velocity is given by $2\varepsilon s$ (where $\varepsilon$ is the hopping amplitude) and, using the relations given in Table 1, the lattice spacing cancels out of the variables in Equation (120). Then for $\bar{U}_\perp = 0 = \bar{W}_\perp$, since $H_s$ and $H_c$ are quadratic forms in boson variables, the energy levels are linear in the Fermi velocities and hence are proportional to $s$. In order to define the continuum limit, this factor of $s$ must be absorbed either into the time and temperature (which multiply $H$ in the partition function and correlation functions) or into all of the coupling constants. These alternatives are equivalent since one of the coupling constants may be chosen as a unit of energy. On the other hand it will be seen that the continuum limit in general requires that $\bar{U}_\perp/2\alpha$ and $\bar{W}_\perp/2\alpha$ vary as powers of $s$, since the energy levels are not linear in these quantities. Using Equations (120) and bearing in mind the factors of $\alpha$ (which are proportional to $s$), this implies that $U_\perp \to 0$ and $W_\perp \to 0$ as $s \to 0$, which does not correspond to the initial extended Hubbard model or $g$-ology picture, as can be seen from Table 1. It would make no difference if the results were valid for small but finite $s$, which would allow $U_\perp$ to be set equal to $U_\parallel$, for the analytical form of the energy scale will turn out to correspond to $|U_\perp/U_\parallel| \ll 1$ and $|W_\perp/W_\parallel| \ll 1$. It is essential to bear this in mind when making comparisons between the results of continuum theories and those of other approaches which retain a discrete lattice.

The problem has been reduced to two sets of independent fermions, representing the collective modes of the system. This transition back and forth from fermions to bosons is a characteristic feature of one-dimensional systems, and since the fermion operators are nonlinear functions of the original boson collective mode variables, even a fairly simple solution in the new representation constitutes a quite elaborate solution of the original problem. This feature is made possible by the kinematic constraints of one-dimensional motion, but, at the same time, it is demanded by the strong correlations which occur for the same reason.

### 4.3. *Solution of the Spinless Fermion Problems*

The arguments will be given in detail for the spin-wave Hamiltonian and the results for the charge degrees of freedom may then be obtained by substitution.

Two approaches have been used to obtain exact solutions. The first, introduced by Luther and Emery,[16] is to choose a particular value of $\bar{U}_\parallel$:

$$\bar{U}_\parallel = -6/5 \tag{121}$$

in order to remove the second term from $H_s$ and make it a free fermion Hamiltonian, which is solvable by Fourier transformation, using Equations



(25) and (36) to obtain

$$\frac{H_s}{v_s} = \frac{4}{5}\sum_k k(a_{1sk}^\dagger a_{1sk} - a_{2sk}^\dagger a_{2sk}) + \frac{U_\perp}{2\alpha}\sum_k (a_{2sk}^\dagger a_{1sk} + a_{1sk}^\dagger a_{2sk}) \tag{122}$$

Then the canonical transformation

$$\begin{aligned} a_{1sk} &= \cos\theta_k \tilde{a}_{1sk} - \sin\theta_k \tilde{a}_{2sk} \\ a_{2sk} &= \sin\theta_k \tilde{a}_{1sk} + \cos\theta_k \tilde{a}_{2sk} \end{aligned} \tag{123}$$

leads to

$$\frac{\tilde{H}_s}{v_s} = \frac{4}{5}\sum_k E_k(\tilde{a}_{1sk}^\dagger \tilde{a}_{1sk} - \tilde{a}_{2sk}^\dagger \tilde{a}_{2sk}) \tag{124}$$

with

$$E_k = \left[k^2 + \left(\frac{5\bar{U}_\perp}{8\alpha}\right)^2\right]^{1/2} \mathrm{sgn}(k) \tag{125}$$

when $\theta_k$ satisfies

$$\tan 2\theta_k = 5\bar{U}_\perp/8k\alpha \tag{126}$$

The transformation (27) should be made to ensure that, at $T > 0$, all negative energy states are full and all positive energy states are empty. Using Equations (124) and (125), it can be seen that, now, there is a gap $\Delta \equiv v_s|\bar{U}_\perp|/2\alpha$ in the spin-wave spectrum, and this will have important consequences for the correlation functions. A finite gap requires $\bar{U}_\perp$ proportional to $s$ in order to cancel the factor $\alpha$, which pushes us to the anisotropic limit $U_\perp/U_\parallel = 0$ as $s \to 0$. This implies that the energy gap of the lattice model must vanish, which is quite usual when taking the continuum limit for, otherwise, the correlation functions will decay exponentially along the lattice and shrink to nothing in the continuum.

The charge-density Hamiltonian $H_c$ may be diagonalized in exactly the same way[24-26] except that $\bar{U}_\parallel$ is replaced by $(-\bar{W}_\parallel)$ in the solubility condition (121), $\bar{U}_\perp$ is replaced by $\bar{W}_\perp$ in Equations (125) and (126), and the factors $\exp(\pm 4ik_F x)$ in Equation (119) prevent the gap from appearing in the low-lying levels unless there is a half-filled band ($k_F = 0$).

These results agree qualitatively with the calculation of Section 2.1. States with one or more spinless fermions correspond to excitations into the separated bands which occur for strong coupling, and, using Table 1, it can be seen that the energy gap appears in the spin excitations when $(U - 2V) < 0$ (attractive coupling if $2|V| < |U|$) or in the charge-density modes when $(U + 6V) > 0$ (repulsive coupling). With the right combination of on-site and near-neighbor coupling, these conditions can be satisfied simultaneously to produce a gap in all collective modes. A comparison of correlation functions will be made later.

Another method of obtaining the energy spectrum for $\bar{U}_\parallel < 0$ or $\bar{W}_\parallel > 0$ has been given by Luther,[27] who used the fact that $H_s$ and $H_c$ are the continuum limits of lattice-fermion Hamiltonians, which are related to a spin-chain model, whose energy spectrum has been found by Johnson *et al.*[28] The spin Hamiltonian is given by Equation (101) and, in the usual way, the Jordon–Wigner transformation

$$S_m^+ = \psi_m^\dagger \exp\left(i\pi \sum_{n=0}^{m-1} \psi_n^\dagger \psi_n\right)$$
$$S_m^z = \psi_m^\dagger \psi_m - \frac{1}{2} \tag{127}$$

may be used to rewrite it in terms of fermion variables $\psi_m$. Then, taking the continuum limit as described in Section 2 [except that $\bar{\psi}_2$ in Equation (18) is replaced by $-i\bar{\psi}_2^\dagger$] leads to Equation (118) with

$$\frac{v_s}{8}(10 + 3\bar{U}_\parallel) = \frac{s}{2}\left(J_x + J_y + \frac{2J_z}{\pi}\right)$$
$$\frac{\pi v_s}{\Delta}(6 + 5\bar{U}_\parallel) = 4sJ_z \tag{128}$$
$$\frac{v_s \bar{U}_\perp}{\alpha} = J_x - J_y$$

The "free" solutions of the spin chain, analogous to a particle–hole excitation of $H_s$, have minimum energy[28]

$$\Delta = 4\pi\left(\frac{\sin\mu}{\mu}\right)|J_x|\left(\frac{l}{4}\right)^{\pi/\mu} \tag{129}$$

with

$$l^2 = \frac{J_x^2 - J_y^2}{J_x^2 + J_z^2} \tag{130}$$

and

$$\mu/\pi = 1 - \theta \tag{131}$$

where $\theta$ is defined in Equation (107). The equation for $\Delta$ is quoted for small $l$, since the requirement $\Delta \sim s$ may be satisfied by choosing $l \sim s^{\mu/\pi}$. The details of the calculation may be found in Luther's paper.[27] To be consistent with the canonical transformations carried out in obtaining $H_s$ and $H_c$, it is actually necessary to use the continuum form[20] of $\theta$:

$$2\theta = \left(\frac{J_x + J_y - 2J_z/\pi}{J_x + J_y + 6J_z/\pi}\right)^{1/2} \tag{132}$$



since exponents depend upon the details of the model. This illustrates once again the difference between taking the continuum limit in the Hamiltonian and in the equations of motion. According to Equations (121) and (128), the solvable value of $\bar{U}_\parallel$ corresponds to the $x-y$ model, $J_z = 0$, and, at this point, Equations (128)–(132) show that $\pi/\mu = 2$ and $\Delta = v_s\bar{U}_\perp/\alpha$, or twice the value obtained from Equation (125), as it should be. For other values of $\bar{U}_\parallel$, Equations (128) and (132) give $\theta = \theta_s$, where

$$\theta_s = \left(\frac{2 + \bar{U}_\parallel}{2 - \bar{U}_\parallel}\right)^{1/2} \tag{133}$$

and, by Equations (129)–(131), $\Delta \sim \bar{U}_\perp^{1/2(1-\theta)}$, which gives the entire dependence of the gap upon $\bar{U}_\perp$. This result has previously been derived[24] from an exact homogeneity relation of the partition function, which will be proved in Section 4.5. When $|\bar{U}_\parallel| \ll 1$, the scale becomes $\Delta \sim \bar{U}_\perp^{-1/\bar{U}_\parallel}$, and this result as well as the remaining dependence upon $\bar{U}_\parallel$ may also be derived from the renormalization group method as described in Section 5.

A qualitatively different effect appears when $\theta \le 1/2$. The spin waves form bound states of energy

$$\varepsilon_n = \Delta \sin\left(\frac{n\pi}{2}\right)\left(\frac{\theta}{1-\theta}\right) \tag{134}$$

with $n = 0, 1, 2, \ldots, n_0 \le \theta^{-1} - 1$, implying that the spinless fermions form bound states in this region, which corresponds to $\bar{U}_\parallel \le -6/5$. Since the continuum theory is Lorentz invariant, the momentum dependence of the excitations is given by $(\varepsilon_n^2 + k^2)^{1/2}$ as in Equation (125) and this enables us to obtain the free energy.

### 4.4. Correlation Functions

The principal difficulty in evaluating the correlation functions of the model is that the various transformations used to diagonalize the Hamiltonian do not lead to expressions that may readily be evaluated. This is in fact quite usual; however, in this case, it is possible to exploit the existence of energy gaps to evaluate correlation functions for $T \gg \Delta$ and for $T \ll \Delta$, and this is sufficient for many purposes.

When $T \gg \Delta$, the backward and umklapp scattering contributions to $H_s$ and $H_c$ may be ignored and the Hamiltonian reduces to the Luttinger model form $H_{LM}$ of Equation (109), whose correlation functions may be evaluated by the method of Section 2, with the difference that the fermions now have a spin. The correlation functions are of the form

$$\chi_\lambda^R = -i\theta(t)\langle[\hat{O}_\lambda^\dagger(x,t), \hat{O}_\lambda]\rangle \tag{135}$$

where the charge-density wave response is generated by the operator

$$\hat{O}_{\mathrm{CDW}}(x)\equiv[\psi^{\dagger}_{2\uparrow}(x)\psi_{1\uparrow}(x)+\psi^{\dagger}_{2\downarrow}(x)\psi_{1\downarrow}(x)]/2^{1/2} \tag{136}$$

the transverse spin-density response by

$$\hat{O}_{\mathrm{SDW}}(x)\equiv\psi^{\dagger}_{2\sigma}(x)\psi_{1,-\sigma}(x) \tag{137}$$

singlet pairing by

$$\hat{O}_{\mathrm{SP}}(x)\equiv\psi^{\dagger}_{2\sigma}(x)\psi^{\dagger}_{1,-\sigma}(x) \tag{138}$$

and triplet pairing by

$$\hat{O}_{\mathrm{TP}}(x)\equiv\psi^{\dagger}_{2\sigma}(x)\psi^{\dagger}_{1\sigma}(x) \tag{139}$$

Using the representation (53) and (54) (with an added spin label) and transforming to charge- and spin-density waves with the aid of Equation (108), the correlation functions separate and may be evaluated by following the route that led to Equations (81) and (82) to find

$$\chi_{\lambda}=\frac{\theta(t)}{2\pi^2 r}\,\mathrm{Im}\{[r^2K(x-v_c't,r)K^*(x+v_c't,r)]^{\theta_c^{\pm1}/2}\times[rK(x-v_s't,r)K^*(x+v_s't,r)]^{\theta_s^{\pm1}/2}\} \tag{140}$$

where $v_c'$ and $\theta_c$ are obtained by setting $V_1=0$, $V_2=W_{\parallel}/2$, and $v_F=v_c$ in Equations (47), (74), and (83) and similarly for $v_s'$, $\theta_s$:

$$\theta_c=\left(\frac{2-\bar{W}_{\parallel}}{2+\bar{W}_{\parallel}}\right)^{1/2} \tag{141}$$

$$\theta_s=\left(\frac{2+\bar{U}_{\parallel}}{2-\bar{U}_{\parallel}}\right)^{1/2} \tag{142}$$

with $\bar{U}_{\parallel}$ and $\bar{W}_{\parallel}$ defined in Equations (120). As in Equation (84), the Fourier transform of $\chi_{\lambda}$ is a power law at $T=0$, and the overall exponents corresponding to the operators in Equations (136)–(139) are $\mu_{\lambda}$, where

$$\left.\begin{aligned}\mu_{\mathrm{CDW}}&=\theta_c+\theta_s-2\\ \mu_{\mathrm{SDW}}&=\theta_c+\theta_s^{-1}-2\\ \mu_{\mathrm{SP}}&=\theta_c^{-1}+\theta_s-2\\ \mu_{\mathrm{TP}}&=\theta_c^{-1}+\theta_s^{-1}-2\end{aligned}\right\}\ T\gg\Delta \tag{143}$$

which specifies the appropriate combinations of $\theta_c^{\pm1}$ and $\theta_s^{\pm1}$ for Equation (140). Once again the phase factors removed in Equation (21) imply that the wave vector $q$ is measured relative to $2k_F$ for the CDW and SDW correlation functions. The factor $\frac{1}{2}$ in the exponents (140) and the plus-or-minus signs in Equation (140) are consequences of the factor $2^{-1/2}$ and the



minus sign in the transformation (108). Thus, at high temperatures, the correlation functions have power law singularities, similar to those found for spinless fermions.

The low-temperature properties are significantly affected by the existence of an energy gap. As in the strong coupling limit, it is plausible that a spin-wave gap prevents a singularity in the response to a perturbation which flips a spin, whereas for a perturbation that treats the spins symmetrically there is a power law singularity determined solely by the charge-density degrees of freedom.[16,29] To see that this is so, Equation (135) is written in terms of the exact eigenstates $|m\rangle$ and eigenvalues $E_m$ of the full Hamiltonian:

$$\chi^R_{\lambda}=\frac{-i\theta(t)}{Z}\sum_{m,n}e^{-\beta E_m+it(E_m-E_n)}\langle m|\hat{O}_{\lambda}(x)|n\rangle\langle n|\hat{O}_{\lambda}|m\rangle+\mathrm{c.c.} \tag{144}$$

where $Z$ is the partition function. The states separate into a charge-density part together with a spin-density part whose contribution to $|m\rangle$ in Equation (144) is forced by the factor $\exp(-\beta E_m)$ to be the ground state at low temperatures, when there is an energy gap. The crucial question now is the existence of a finite ground-state expectation value for the spin-wave part of $\hat{O}_{\lambda}$, because transitions to excited states will have a minimum energy $\Delta$ and give an oscillating factor $\exp(i\Delta t)$ which will prevent the divergence of the Fourier transform of $\chi^R_{\lambda}$ at low frequencies. This can be examined[29] most easily for the solvable case $\bar{U}_{\parallel}=-6/5$. In the boson representation, the spin-wave parts of $\hat{O}_{\mathrm{SDW}}$ and $\hat{O}_{\mathrm{TP}}$ are proportional to $\exp[2^{-1/2}i(\Phi_{1s}+\Phi_{2s})]$, whereas $\hat{O}_{\mathrm{CDW}}$ and $\hat{O}_{\mathrm{SP}}$ contain $\exp[2^{-1/2}i(\Phi_{1s}-\Phi_{2s})]$. The transformation $\exp(iS)$ used in Equation (114) removes the factor $2^{-1/2}$ in the first of these factors and gives the boson representation of $\psi_{1s}\psi_{2s}$, which has a zero ground-state expectation value because it does not conserve particle number. The second factor (related to $\hat{O}_{\mathrm{CDW}}$ and $\hat{O}_{\mathrm{SP}}$) becomes $\exp[i(\Phi_{1s}-\Phi_{2s})/2]$, which is (crudely speaking) proportional to $(\psi^{\dagger}_{2s}\psi_{1s})^{1/2}$ and has a finite ground-state expectation value. In that case the spin degrees of freedom are frozen and do not contribute to the power law singularity, but they do not inhibit the charge-density waves. A similar argument gives the consequences of a charge-density wave gap for $W_{\parallel}>0$, and a half-filled band, which removes the singularity in $\chi^R_{\mathrm{SP}}$ and $\chi^R_{\mathrm{TP}}$.

The overall conclusions for $\bar{U}_{\parallel}<0$ are then as follows.

### 4.4.1. Band Not Half-Filled

Here the umklapp scattering is ineffective because $\bar{k}_F\neq0$ and the contribution from the charge-density waves is the same as at high

temperatures. Thus, $\chi^R_{SDW}$ and $\chi^R_{SP}$ are nonsingular whereas

$$\left.\begin{aligned}\mu_{CDW} &= \theta_c \\ \mu_{SP} &= \theta_c^{-1}\end{aligned}\right\} \quad T \ll \Delta \tag{145}$$

### 4.4.2. Half-Filled Band, $\bar{W}_\parallel > 0$

Here, $\chi^R_{SDW}$, $\chi^R_{SP}$, and $\chi^R_{TP}$ have no divergence but $\chi^R_{CDW}$ is a constant at large $x$ and $t$ since both ground-state matrix elements are finite. This implies that the ground state has long-range CDW order even for the one-dimensional system, a consequence of the inhibition of quantum fluctuations by the energy gaps.

Although this discussion was restricted to the solvable values of $\bar{U}_\parallel$ and $\bar{W}_\parallel$, it is plausible that the energy gaps established for other values of $\bar{U}_\parallel (<0)$ and $\bar{W}_\parallel (>0)$ have the same consequences, particularly as they are in qualitative agreement with the strong coupling results.

The picture is not yet complete because the methods used so far tell us nothing about solutions for $\bar{U}_\parallel > 0$ or $\bar{W}_\parallel < 0$ (unless umklapp scattering is ineffective). The solvable values of the coupling constants have the opposite signs and the spin-chain equivalence may not be used because, as will be seen in Section 4.5.1, the continuum limit does not exist, at least as it has been defined so far. The renormalization group method described in Section 5 will enable us to fill in the missing information, and a further discussion of the phases and correlation functions will be deferred until then.

## 4.5. Relationship to Other Problems

At this point the development of the theory will be interrupted in order to discuss the relation between the spinless fermion Hamiltonians and a number of other problems—the sine–Gordon equation, classical Coulomb gas, $x$–$y$ model, and interface roughening.

### 4.5.1. Sine–Gordon Equation

Suppose the $H_{LM}$ in Equation (109) is diagonalized by canonical transformation, as in Section 2. Then the boson form of the backward and umklapp scattering terms (110) and (111) may be rewritten (for $\bar{k}_F = 0$)

$$H_{BS} = \frac{s\bar{U}_\perp}{2\pi^2\alpha}\int dx \cos[(8\pi\theta_s)^{1/2}\Phi_s(x)] \tag{146}$$

$$H_u = \frac{s\bar{W}_\perp}{2\pi^2\alpha}\int dx \cos[(8\pi\theta_c)^{1/2}\Phi_c] \tag{147}$$



where $\theta_s$, $\theta_c$ are defined in Equations (141) and (142) and, using Equations (112) and (113), $\Phi_s = (\Phi_{1s} - \Phi_{2s})/(4\pi)^{1/2}$, $\Phi_c = (\Phi_{1c} - \Phi_{2c})/(4\pi)^{1/2}$ are canonical boson fields. [The factors $(4\pi)^{1/2}$ come from comparison of Equation (54) with the standard definition of a Bose field.] In this form, $H_s$ and $H_c$ are sums of a free boson term and the nonlinear cosine contributions of Equations (146) or (147). They constitute sine–Gordon Hamiltonians, and the relationship with results that were obtained independently and almost simultaneously in elementary particle physics becomes evident. Coleman[30] started from the sine–Gordon equation and used perturbation theory in $\bar{U}_\perp$ or $\bar{W}_\perp$ to derive the relationship to spinless fermions. This stimulated Mandelstam[31] to rediscover the boson representation of fermion operators as a simpler method of carrying out the calculation! (See also the work of Mattis.[32])

The connection between these two lines of development was noticed by Peschel[33] and by Heidenreich *et al.*,[34] and it had a number of immediate implications. The sine–Gordon equation had been studied extensively in both classical and quantum mechanics, and it was known to have particlelike solutions called solitons, which were just the spinless fermions of Section 4 and had a mass equal to the energy gap. The existence of soliton bound states and the expression (134) for their spectrum had been derived by a semiclassical approximation.[35] It was clear that much of the behavior of one-dimensional systems could be regarded as soliton processes, and reviews of this point of view have been given by Emery[3] and by Luther.[36]

One further result, which will be useful later and is most easily proved[30] by applying the variational principle to the sine–Gordon equation, is that the ground state is unbounded below for $\bar{U}_\parallel > 0$ or $\bar{W}_\parallel < 0$ (i.e., $\theta_s > 1$, $\theta_c > 1$) when the continuum limit is taken in the manner described in Section 4.3. This implies that the connection to the spin-chain problem breaks down in this region, since it relies on the continuum limit, and indeed, at present, we do not know the precise range of values of $\theta_s$ or $\theta_c$ for which the whole procedure is well defined. The same difficulty does not occur if there is a finite bandwidth (as already known for the analogous problem of magnetic impurities in metals[37]) but then it is necessary to ascertain which of the results proved in the continuum limit are still correct, or to use a quite different approach. This question will be taken up in Section 5.

### 4.5.2. Classical Coulomb Gas

The equivalence to the classical Coulomb gas[24,38] may be obtained by expanding the partition functions $Z_s$ for the spin-density waves in powers of $U_\perp$. Every order is an integral of the expectation value of a power of

$\cos[(8\pi\theta_s)^{1/2}\Phi_s]$ and therefore is a correlation function of Luttinger's model. Then, using the method of Section 2, it is not difficult to obtain

$$Z_s = Z_0 \sum_{m=0}^{\infty} \frac{1}{(m!)^2}\left[\frac{\bar{U}_\perp}{4\pi\alpha^2}\right]^{2m} \int_0^L \prod_i^{2m} dx_i \int_0^{v_s'\beta} \prod_j^{2m} dv_s' t_j \times \exp\left(\sum_{p,q}^{2m} (-1)^{p-q} C_{pq}\right) \tag{148}$$

where $v_s'$ is the transformed Fermi velocity, $Z_0$ is the partition function for $U_\perp = 0$ and $C_{pq} \equiv C(x_p - x_q, t_p - t_q)$ with

$$e^{-C(x,t)} = |\bar{K}(x - iv_s' t, \alpha)\bar{K}(-x - iv_s' t, \alpha)|^{\theta_s} \tag{149}$$

and $\bar{K}(\xi, \alpha)$ given by Equation (68). Equation (148) may now be given a quite different interpretation by regarding $v_s' t$ as a distance and $v_s'\beta$ as the total length in that direction. Then $Z_s$ is the grand partition function of a *classical* system with fugacity proportional to $U_\perp$. The thermodynamic limit requires $\beta \to \infty$ when, using Equations (64) and (149),

$$C(x, t) = \theta_s \ln\{[x^2 + (v_s' t)^2]/\alpha^2\} \tag{150}$$

becomes the Coulomb potential in two dimensions. Thus the ground-state properties of the quantum system are related to the thermodynamic behavior of the classical problem, and variation of the coupling constants corresponds to a change in fugacity and temperature. Since the thermal average of an odd number of fermion fields vanishes, only even powers of $U_\perp$ occur and the Coulomb gas is neutral. There is a similar series for $Z_c$.

This equivalence gives an appealing picture of the energy gap as a screening length,[38] but it adds little to our mathematical understanding of the initial quantum system since much of what is known about the classical Coulomb gas in two dimensions comes from renormalization group equations[39] that are one order lower than those to be obtained in Section 5. At the same time, the problem is sufficiently delicate for intuition developed from experience with the three-dimensional gas to be untrustworthy. Nevertheless it is possible to give a simple derivation of the scaling relations of the continuum limit, and, in fact, they were first obtained in this way.[24,37] They are a consequence of the fact that $Z_s$ is a homogeneous function of degree zero in $\beta$, $L$, and $\alpha$ since, if these quantities were divided by a factor $\lambda$, it would be possible to recover the original form of $Z_s$ by changing integration variables to $x_i' = \lambda x_i$ and $t_i' = \lambda t_i$. Further, for small $\alpha$, all factors $\alpha$ and $\bar{U}_\perp$ appear in the combination $\bar{U}_\perp \alpha^{2(\theta_s - 1)}$. It follows that the continuum limit may be taken by absorbing $s$ into $t_1$ while requiring $\bar{U}_\perp \sim \alpha^{2(1-\theta_s)}$, and that, since the energy gap $\Delta$ is an intensive quantity and $Z_s$ is a function of $\beta\Delta$, homogeneity implies $\Delta \sim \alpha^{-1}$ or, more completely, $\Delta \sim \alpha^{-1} \bar{U}_\perp^{1/2(1-\theta_s)}$. These results are in complete agreement with Equations



(128)–(131) and the definition of the continuum limit given in Section 4.4.

Although the argument presented here relied on taking $s \to 0$, it is also possible to relate the fermion system with finite bandwidth to a classical gas, but one with complicated many-body interactions, except at large distances where they separate into two-body Coulomb forces. To the extent that the long-range parts of the interactions determine the asymptotic behavior of correlation functions, it might be expected that general properties do not depend sensitively upon the magnitude of the bandwidth, but this does not necessarily mean that the continuum limit is harmless, and it is necessary to be aware that taking the bandwidth to infinity can have significant effects. This will become still more evident in Section 5.

### 4.5.3. *Interface Roughening and the Classical x–y Model*

Finally it is of interest to mention that the Coulomb gas is equivalent to the discrete Gaussian model of interface roughening[40] for arbitrary dimension, and dual to Villain's form of the classical $x$–$y$ model in two dimensions.[41–43] The rough phase of the former[40–44] and the low-temperature phase of the latter[39,41] model constitute a line of critical points and have analogs in the one-dimensional quantum system which will be explored in Section 5.

It is important to keep in mind these relationships between completely different physical systems, since progress in any one field is likely to have a substantial impact on all of the others.

## 5. *Renormalization Group Method*

The methods described in Section 4 have enabled us to obtain solutions for $U_\parallel < 0$ and for arbitrary $W_\parallel$ unless there is a half-filled band, when they are restricted to $W_\parallel > 0$. In order to deal with other couplings and to get a better appreciation of the significance of the continuum limit, it is necessary to turn to some other approach, and in this section it will be shown that the renormalization group technique, to a great extent, is able to fill in the missing information, although by itself it is not able to provide a complete solution. The idea is to keep a finite bandwidth or lattice spacing and to vary it while making a compensating change of the coupling constants so as to leave the partition function unaffected. In this way, one physical problem may be scaled step-by-step onto another, which may be easier to solve. The most complete discussions of the scaling equations and their consequences have been given by Menyhárd and Sólyom[45] and by Sólyom,[46] but their approach will not be followed here because it is possible to derive more general equations by a simpler method and also to

find more accurate solutions, making use of results that have been obtained in the preceding sections.

### 5.1. Scaling Equations

The consequences of varying the lattice spacing in Equation (148) with $C(x, t)$ given by Equation (150), may be obtained at once from the scaling law $\tilde{U}_\perp \sim \alpha^{2(1-\theta_s)}$ derived in Section 4.5. On defining $l \equiv \ln \alpha$ and differentiating, it is found that

$$\frac{d\tilde{U}_\perp}{dl} = 2(1-\theta_s)\tilde{U}_\perp \tag{151}$$

where the notation $\tilde{U}_\perp$ has been used for the scaled value of $\tilde{U}_\perp$.

This is a complete renormalization group equation for Equation (148) in which $\alpha$ has been set equal to zero wherever possible, but it certainly is not the whole story for the original Fermi system with finite bandwidth. The right-hand side of Equation (151) is linear in $\tilde{U}_\perp$ because a variation of $\alpha$ has produced a mere multiplicative change in the correlation functions which may be absorbed order by order, and the coefficient is correct because all such changes must be compensated by choice of $\tilde{U}_\perp$ if a well-defined continuum limit is to be obtained. When the omitted dependence on lattice spacing is restored, a variation in $s$ in one order must be compensated by adjusting $\tilde{U}_\perp$ in a different order, and this adds higher powers of $\tilde{U}_\perp$ to the right of Equation (151) as well as generating a similar equation for $\tilde{U}_\parallel$.

For present purposes, it is sufficient to know the scaling equations to third order in $\tilde{U}_\parallel$ and $\tilde{U}_\perp$, and they may be obtained from Equation (151) if it is supplemented by simple symmetry arguments. First of all, the sign of $\tilde{U}_\perp$ in $H_s$ [Equation (118)] may be reversed by a gauge transformation which has no physical effect and should therefore leave the scaling equations unchanged. Then, to third order, using Equation (133) to expand $\theta_s$ in powers of $\tilde{U}_\parallel$,

$$\frac{d\tilde{U}_\perp}{dl} = -\tilde{U}_\perp\left(\tilde{U}_\parallel + \frac{\tilde{U}_\parallel^2}{4} + c\tilde{U}_\perp^2\right) \tag{152}$$

Now, if Equation (109) were rewritten in terms of the original fermion variables, the $U_\parallel$ term would differ from $H_{BS}$ in Equation (110) only in that all fields would have the same spin and, wherever both interactions play the same role, they will give the same numerical coefficients, so $c = 1/4$ in Equation (152). The equation for $\tilde{U}_\parallel$ may be written down immediately from the requirements that (i) it is even in $\tilde{U}_\perp$, (ii) there is no renormalization when $\tilde{U}_\perp = 0$, and (iii) it becomes identical to the equation



for $\tilde{U}_\perp$ when $\tilde{U}_\parallel = \tilde{U}_\perp$. The latter follows because spin rotational invariance implies that, physically, $\tilde{U}_\parallel = \tilde{U}_\perp$ (see Table 1) and this symmetry is not spoiled by renormalization. Thus, to third order, the scaling equations are

$$\frac{d\tilde{U}_\perp}{dl} = -\tilde{U}_\perp\left(\tilde{U}_\parallel + \frac{\tilde{U}_\parallel^2 + \tilde{U}_\perp^2}{4}\right) \tag{153}$$

$$\frac{d\tilde{U}_\parallel}{dl} = -\tilde{U}_\perp^2\left(1 + \frac{\tilde{U}_\parallel}{2}\right) \tag{154}$$

In the usual way, the equivalent equations for the charge-density waves may be obtained by substituting $(-\tilde{W}_\parallel, \tilde{W}_\perp)$ for $(\tilde{U}_\parallel, \tilde{U}_\perp)$, although there is renormalization only for a half-filled band. The advantage of the notation is apparent for, not only does it allow an enormous simplification of the derivation, but also the separation of the scaling equations into independent sets comes about quite naturally, whereas, in other definitions,[45–47] this is apt not to be the case.

### 5.2. Trajectories and Energy Scales

The scaling trajectories may be obtained by eliminating $l$ from Equations (153) and (154) and solving to obtain

$$\frac{\tilde{U}_\parallel^2 - \tilde{U}_\perp^2}{2 + \tilde{U}_\parallel} = a \tag{155}$$

where $a$ is a constant. This equation describes families of hyperbolas displaced along the $\tilde{U}_\parallel$ axis, as illustrated in Figure 3. The low-order scaling equations may be used to solve the problem in region I, in which $\tilde{U}_\parallel \geq |\tilde{U}_\perp|$ for, according to Equation (153), $d|\tilde{U}_\perp|/dl \leq 0$, so $|\tilde{U}_\perp|$ will continue to decrease until it meets the "fixed line"

$$\tilde{U}_\perp = 0. \tag{156}$$

for which the right-hand sides of Equations (153) and (154) vanish and

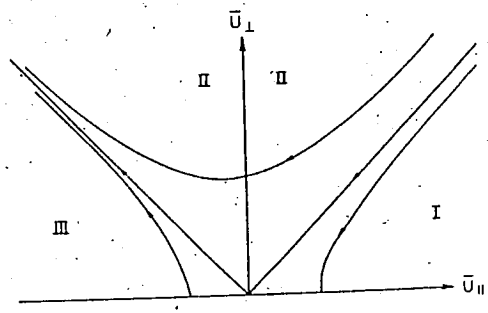


Figure 3. Scaling trajectories in the $(\tilde{U}_\parallel, \tilde{U}_\perp)$ plane. In region I $(\tilde{U}_\parallel \geq |\tilde{U}_\perp|)$ the renormalization stops at the fixed line $\tilde{U}_\perp = 0$ whereas in regions II and III it goes out to strong coupling.

renormalization stops. The asymptotic value of $\tilde{U}_\parallel$ is $\tilde{U}_R$ where, substituting the initial and final coupling constants into Equation (155)

$$\frac{\tilde{U}_R^2}{2+\tilde{U}_R}=\frac{\tilde{U}_\parallel^2-\tilde{U}_\perp^2}{2+\tilde{U}_\parallel} \tag{157}$$

In this case, an initially weak coupling problem scales to still weaker coupling, so the trajectories stay within the region of validity of Equations (153) and (154), and the asymptotic Hamiltonian is a Luttinger model ($\tilde{U}_\parallel=\tilde{U}_R$, $\tilde{U}_\perp=0$), which is solvable. The correlation functions and their physical consequences will be given in Section 5.3.

The fixed line (156) is a property of the exact scaling equations, since there is no renormalization when $\tilde{U}_\perp=0$, and it corresponds to the critical line common to the $x$–$y$ model[39,41] and the interface roughening models,[40,44] for which the correlation functions are power laws and there is no long-range order.

On the other hand, in regions II and III of Figure 3, $\tilde{U}_\parallel$ and $\tilde{U}_\perp$ scale towards strong coupling and the behavior is governed by the "fixed point"

$$|\tilde{U}_\perp|=\tilde{U}_\parallel =2 \tag{158}$$

but this is outside the range of validity of the scaling equations. There is a fixed point at infinitely strong coupling (corresponding to infinite temperature in the Coulomb gas) but the approximation has placed it at the finite value (158) giving the impression of a critical fixed point. The actual "phase transition," or change in ground state of the quantum system, takes place when the coupling constant values cross the line $\tilde{U}_\parallel=|\tilde{U}_\perp|$. This is strongly suggested by the general features of Figure 3, with one state corresponding to a flow to strong coupling and the other to a flow towards $\tilde{U}_\perp=0$, and it is completely consistent with the results of Sections 3 and 4, where it was found that, for particular locations in regions II and III of Figure 3, there is an energy gap that is essentially the finite coherence length of the disordered phase. Specifically, this was shown for the continuum limit when $|U_\perp|\ll -U_\parallel$ and for strong coupling when $U_\parallel=U_\perp<0$. Also, at the end of Section 3, it was pointed out that, in the strong coupling limit with $U<0$, a gap appears in the charge-density wave spectrum when $V$ passes through 0 and, using Table 1, this is the same as crossing the isotropic line from region I to region II in the $(-\tilde{W}_\parallel\tilde{W}_\perp)$ plane.

Although it is necessary to consider all values of $\tilde{W}_\parallel$ and $\tilde{W}_\perp$, the situation is somewhat simpler for the spin-density waves because, physically, $\tilde{U}_\parallel=\tilde{U}_\perp$, so that scaling proceeds along the isotropic lines $\tilde{U}_\parallel=\pm|\tilde{U}_\perp|$ and the transition takes place at $\tilde{U}_\parallel=0$. Also, since the continuum limit requires that $\tilde{U}_\perp\to 0$ and $\tilde{W}_\perp\to 0$, it will give a transition at



$\tilde{U}_\parallel=0$ or $\tilde{W}_\parallel=0$, which is consistent with the explicit solution found in Section 4.

There remains the problem of finding a solution for an arbitrary point in regions II and III. One possibility is to scale onto the strong coupling limit and then to use the method of Section 3, but this must be done numerically. Another that has been considered[16] is to scale onto the solvable value $\tilde{U}_\parallel=-6/5$, but this gives no more than a qualitative indication of the existence of a gap for several reasons—Equations (153) and (154) are not accurate at such a large value of $|\tilde{U}_\parallel|$, the scaling procedure generates many-body forces that are not likely to be negligible, and, finally, $\tilde{U}_\perp$ becomes large, which is incompatible with the continuum limit. Nevertheless, all of the evidence points to the existence of energy gaps in regions II and III or effective Luttinger models in region I and, as argued in Section 4.4, this is sufficient to determine the correlation functions at temperatures that are much higher than or much lower than the energy gap.

If this information were supplemented by a determination of the crossover temperature, the picture of the physical behavior of the one-dimensional electron gas would be reasonably complete. At weak coupling, the third-order scaling equations are sufficient for this purpose. Using Equation (155) to eliminate $U_\perp$ from Equation (154) and integrating gives

$$\int_{\tilde{U}_\parallel}^{\tilde{U}_\parallel} dU\,\{(1+U/2)[U^2-a(2+U)]\}^{-1}=\ln(\alpha/\tilde{\alpha}) \tag{159}$$

where $\tilde{\alpha}$ is the scaled value of $\alpha$. If the indefinite integral on the left of this equation is denoted by $\phi(U)$, it can be seen that $\tilde{U}_\parallel$ is a function of $\alpha^{-1}\exp[-\phi(\tilde{U}_\parallel)]$, which gives the dependence of physical quantities upon the original parameters of the problem once they have been evaluated in terms of $\tilde{U}_\parallel$. In other words, it defines the energy scale. The integral can be evaluated in closed form for arbitrary $a$ but there are two situations of particular interest. For the spin-density waves, the physical value is $a=0$, [or $\tilde{U}_\parallel=\tilde{U}_\perp$ in Equation (155)], which leads to $\phi(U)=-U^{-1}-(\ln|U|)/2+O(U)$ and an energy scale

$$E_s=\frac{A}{\alpha}|\tilde{U}_\parallel|^{1/2}e^{\tilde{U}_\parallel^{-1}},\qquad \tilde{U}_\perp=\tilde{U}_\parallel \tag{160}$$

This result is identical[24] to that obtained from exact solutions of the Hubbard model,[48] and a slightly more detailed analysis enables the coefficient $A$ to be obtained.[49] The other interesting region is $|\tilde{U}_\perp|\ll|\tilde{U}_\parallel|\ll 1$, which may be compared to the solutions in the continuum limit. In this case,

$$E_s=\frac{A}{\alpha}|\tilde{U}_\parallel|^{1/2}\left|\frac{\tilde{U}_\perp}{2\tilde{U}_\parallel}\right|^{1/4-\tilde{U}_\parallel^{-1}},\qquad |\tilde{U}_\perp|\ll|\tilde{U}_\parallel| \tag{161}$$

The exponent of $\tilde{U}_\perp$ is the same as the first two terms in the expansion of $(1-\theta_s)^{-1}/2$ and, to this order, it agrees with the form of the continuum limit given by Equations (129) and (131) or by scaling the Coulomb gas free energy. However, the remaining dependence on $|\tilde{U}_\parallel|$ in Equation (161) is not the same as in Equation (129) for small $|\tilde{U}_\parallel|$, and it is clear that taking the continuum limit as prescribed in Section 4 has nontrivial consequences. The higher orders in $|\tilde{U}_\perp|$ which were introduced in going from Equation (151) to Equations (153) and (154) affect the energy scale even when $|\tilde{U}_\perp|$ is small. In considering the significance of this result, it is necessary to bear in mind that it is not known if the continuum limit exists for all negative values of $|\tilde{U}_\parallel|$ and, whenever it does not exist, the spin-chain solution is not directly applicable.

### 5.3. Low-Temperature Properties

We are now in a position to evaluate high- and low-temperature correlation functions for $\tilde{U}_\parallel > 0$ and $\tilde{W}_\parallel < 0$, filling in the gaps that were left at the end of the discussion of the continuum limit. The physical properties may be inferred from the correlation functions, which have the power-law behavior of Luttinger's model, modified by the existence of an energy gap and renormalized exponents. A different approach was used by Sólyom,[46] who wrote down scaling equations for the correlation functions and evaluated them entirely within the renormalization group method, without recourse to the exact solution of Luttinger's model. This has the disadvantage that one is forced to rely on the fixed point (158), which is a fiction and does not reveal the energy gap. Furthermore, the exponents cannot be found to a higher accuracy than the scaling equations, and in this respect are clearly less satisfactory. However, the scaling equations for the correlation functions do expose logarithmic corrections,[46] which usually may be ignored (and will be ignored here) but become important when the exponents vanish.

The correlation functions may be specified by giving their exponents as in Equations (141)–(143) and (145). They are listed in Table 2 for high and low temperatures for a band that is not half-filled, and in Table 3 for low temperatures and a half-filled band. The functions $\theta_{cR}$ and $\theta_{sR}$ are obtained by replacing $\tilde{W}_\parallel$ and $\tilde{U}_\parallel$ by $\tilde{W}_R$ and $\tilde{U}_R$ in Equations (141) and (142), respectively. For physical systems, $\tilde{U}_\parallel = \tilde{U}_\perp$ and $\tilde{U}_R = 0$ so $\theta_{sR} = \theta_{sR}^{-1} = 1$. Where no exponent is quoted, it is implied that the power law singularity has been removed by an energy gap. In this connection, notice that Table 3 has no CDW entry for $\tilde{W}_\parallel > -|\tilde{W}_\perp|$ and $\tilde{U}_\parallel \geq |\tilde{U}_\perp|$, whereas the argument of Section 4 leads to an exponent of $\theta_{sR} - 2$. The reasons for this conclusion are given in Appendix C.



*Table 2. Correlation Function Exponents: Band Not Half-Filled*

| Correlation function | High temperature $T \gg \Delta$ | Low temperature $T \ll \Delta$ | |
|---|---|---|---|
| | | $\tilde{U}_\parallel < \lvert\tilde{U}_\perp\rvert$ | $\tilde{U}_\parallel \geq \lvert\tilde{U}_\perp\rvert$ |
| CDW | $\theta_c + \theta_s - 2$ | $\theta_c - 2$ | $\theta_c + \theta_{sR} - 2$ |
| SDW | $\theta_c + \theta_s^{-1} - 2$ | — | $\theta_c + \theta_{sR}^{-1} - 2$ |
| SP | $\theta_c^{-1} + \theta_s - 2$ | $\theta_c^{-1} - 2$ | $\theta_c^{-1} + \theta_{sR} - 2$ |
| TP | $\theta_c^{-1} + \theta_s^{-1} - 2$ | — | $\theta_c^{-1} + \theta_{sR}^{-1} - 2$ |

When $q = 2k_F$ for the density waves or $q = 0$ for pairing states, the correlation functions vary as $\omega^\mu$, where $\mu$ is tabulated. Negative values of $\mu$ imply that long-range correlations will develop at low temperatures, as shown in Figure 4 for a band that is not half-filled and Figure 5 for a half-filled band. Both figures assume the physical condition $\tilde{U}_\parallel = \tilde{U}_\perp$. Figure 4 uses $\tilde{W}_\parallel$ and $\tilde{U}_\parallel = \tilde{U}_\perp$ as variables but, since $|\tilde{W}_\perp|$ is significant for a half-filled band, it is more appropriate to display $\tilde{W}_\parallel$ and $|\tilde{W}_\perp|\,\mathrm{sgn}(\tilde{U}_\parallel)$ in Figure 5. The "phase diagrams" in terms of the variables of the original models may be obtained with the aid of Table 1.

It is interesting to compare Figure 5 to the conclusions of the strong $U$ limit which, using Table 1, corresponds to the neighborhood of the isotropic lines in the first and third quadrants. Then $U > 0$ gives SDW correlations and $U < 0$ gives CDW correlations if $V > 0$, or a combinations of CDW and SP correlations when $V < 0$. This is in agreement with the discussion at the end of Section 3.

An issue of considerable practical interest is the possibility of making a high-temperature superconductor out of organic metals. Since coupling between chains commonly favors a charge-density wave state which creates a single-particle gap and destroys superconductivity, it is desirable to have a system that inclines to pairing but not CDW correlations. Figure 4 and 5 both show that $\tilde{U}_\parallel > 0$ and $\tilde{W}_\parallel < 0$ is the most favorable case, and according

*Table 3. Correlation Function Exponents for a Half-Filled Band at Low Temperatures*

| Correlation function | $\tilde{W}_\parallel \leq -\lvert\tilde{W}_\perp\rvert$ | | $\tilde{W}_\parallel > -\lvert\tilde{W}_\perp\rvert$ | |
|---|---|---|---|---|
| | $\tilde{U}_\parallel < \lvert\tilde{U}_\perp\rvert$ | $\tilde{U}_\parallel \geq \lvert\tilde{U}_\perp\rvert$ | $\tilde{U}_\parallel < \lvert\tilde{U}_\perp\rvert$ | $\tilde{U}_\parallel \geq \lvert\tilde{U}_\perp\rvert$ |
| CDW | $\theta_{cR} - 2$ | $\theta_{cR} + \theta_{sR} - 2$ | $-2$ | — |
| SDW | — | $\theta_{cR} + \theta_{sR}^{-1} - 2$ | — | $\theta_{sR}^{-1} - 2$ |
| SP | $\theta_{cR}^{-1} - 2$ | $\theta_{cR}^{-1} + \theta_{sR} - 2$ | — | — |
| TP | — | $\theta_{cR}^{-1} + \theta_{sR}^{-1} - 2$ | — | — |

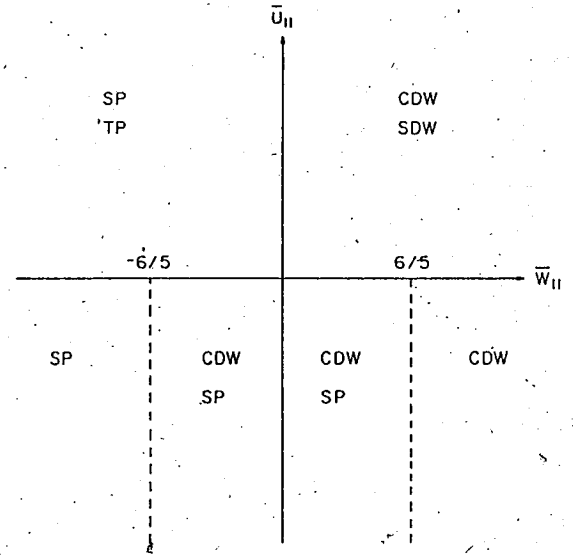


Figure 4. "Phase diagram" in the $(\bar{W}_\parallel, \bar{U}_\parallel)$ plane for a not half-filled band, showing regions in which the various correlation functions diverge as $\omega \to 0$ and $T \to 0$.

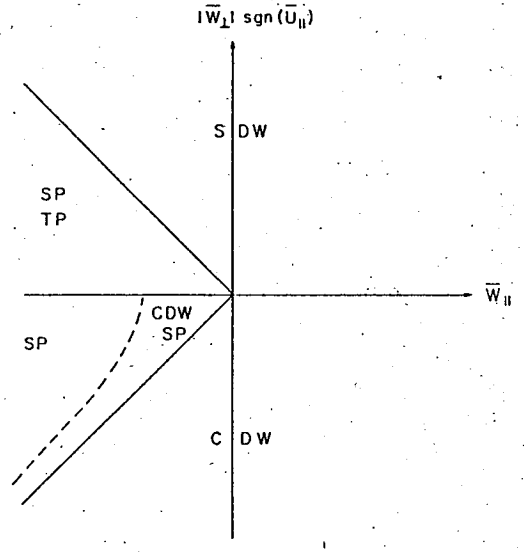


Figure 5. "Phase diagram" in the $(\bar{W}_\parallel, |\bar{W}_\perp| \operatorname{sgn}(\bar{U}_\parallel))$ plane for a half-filled band showing regions in which the various correlation functions diverge as $\omega \to 0$ and $T \to 0$. The dashed line is $\bar{W}_R = -6/5$.



to Table 1, this requires $U > 0$, $V < -U/6$. that is, a short-range repulsion and a longer-range attraction. This interaction, which also is characteristic of superfluid[3] He,[5] may be achieved if the Coulomb force is offset by exchange of phonons or electronic collective modes at long distances, but it does require a fairly strong attraction if $U$ is not reduced by molecular polarization.

### 5.4. Four-Particle Functions

An important feature of the boson representations is that it is not too difficult to evaluate many-particle correlation functions, and the arguments given in Section 2 show that the method is quite generally valid for this purpose. Indeed it has already been used in Section 4.5 to derive the Coulomb gas partition function. Up to this point, the discussion of physical properties has concentrated upon two-particle correlations, but x-ray scattering experiments[50,51] appear to have shown signs of divergences in certain four-particle functions,[52] and therefore it is worthwhile showing how this comes about.

Experimentally, the existence of a charge-density or spin-density wave state is most clearly seen in x-ray or neutron scattering experiments which usually see directly the lattice effects produced by the electron–phonon coupling

$$H_{ep} = \int dx\, \phi(x)[g_2\, e^{2ik_F x}\hat{O}_{CDW} + g_4\, e^{4ik_F x}\hat{O}_4] + \text{H.c.} \tag{162}$$

where $\hat{O}_{CDW}$ is given in Equation (136),

$$\hat{O}_4(x) = \psi^\dagger_{2\uparrow}(x)\psi^\dagger_{2\downarrow}(x)\psi_{1\downarrow}(x)\psi_{1\uparrow}(x) \tag{163}$$

and $\phi(x)$ is the phonon field. The factors $\exp(2ik_F x)$ and $\exp(4ik_F x)$ have been exposed by the transformations (21) and (22). Physically, $H_{ep}$ originates in the modulation of the electron hopping $H_0$ and Coulomb interaction $H_2$ which occurs when a phonon produces a local variation of the lattice spacing. Only large momentum transfers have been retained, in order to focus on the processes in which one or two particles are transferred across the Fermi surface, giving rise to the singular effects.

The electrons and phonons form a coupled system and, in reality, neither one can be treated in isolation from the other, but space does not permit an adequate discussion of this problem, and, for present purposes, Equation (162) is merely introduced to indicate that the lattice couples to two-particle and four-particle operators in such a way that a tendency to form long-range electronic correlations will produce a measurable effect on the lattice. It is clear from Equation (162) that $\hat{O}_{CDW}$ and $\hat{O}_4$ have their singular behavior at $q = 2k_F$ and $q = 4k_F$, respectively, reflecting the

change in momentum when two or four electrons are scattered across the Fermi surface. The momentum balance has to be taken up by the phonons and they will respond at the corresponding wave vectors. The operator $\hat{O}_4$ has already appeared in the umklapp scattering term given in Equation (111), and the boson representation (113) involves only the charge-density wave operators. It is then straightforward to show, as before, that the corresponding correlation function behaves as

$$\chi_4^R \sim \omega^{4\theta_c - 2} \tag{164}$$

for $q = 4k_F$. From the definition (141) of $\theta_c$, it can be seen that $\chi_4^R$ diverges when $\omega \to 0$, provided $\bar{W}_\parallel > 6/5$, which requires a fairly strong repulsive interaction. The low-temperature state which occurs when there are coupled chains is a correlated state of charge-density waves mixed with a lattice distortion, and it is much more complicated than states that normally are envisaged for more isotropic systems.

For a half-filled band, Equation (164) gives only the high-temperature behavior, and the normalization group argument indicates that, at low temperatures, if $\bar{W}_\parallel \leq -|\bar{W}_\perp|$, $\bar{W}_\parallel$ must be replaced by $\bar{W}_R$, otherwise there is a gap that freezes out the charge-density waves and gives an exponent of $(-2)$.

It is also interesting to compare the exponents of $\chi_{\mathrm{CDW}}^R$ and $\chi_4^R$ in order to see which one dominates at any particular temperature. For repulsive interactions and the physical situation $\bar{U}_\parallel = \bar{U}_\perp$, Table 2 shows that $\chi_{\mathrm{CDW}}^R$ has an exponent which crosses over from $\theta_c + \theta_s - 2$ to $\theta_c - 1$ as the temperature decreases, and the CDW response becomes more divergent. At low temperatures, $\chi_{\mathrm{CDW}}^R$ is more divergent than $\chi_4^R$ when $\bar{W}_\parallel < 8/5$, and there can be a crossover from a dominant $4k_F$ response at high temperatures to a dominant $2k_F$ response at low temperatures, provided $\bar{U}_\parallel$ and hence $\theta_s$ is large enough. It is possible that this mechanism[52] accounts for the observed x-ray scattering[50,51] at $4k_F$ observed in TTF-TCNQ and, if so, it indicates that the electron–electron coupling is repulsive and fairly strong.

The transformation (114) changes the four-particle part of $H_{\mathrm{ep}}$ into an interaction between a phonon and *two* spinless fermions and therefore it should be possible to construct essentially the same theory to describe the interplay between the lattice and either kind of correlated state.[52]

There are divergences in other four-particle correlation functions,[53] but, so far, no experimental effects have been found that might be attributable to them.

The dynamical role of phonons is not the only significant topic that has been omitted from the discussion. The production of phase transitions by interchain coupling, properties of the ordered phases, effects of impurities, intramolecular vibrations, and transport properties are important subjects



that have been omitted, although some of them are discussed in other chapters of this volume. I am aware that I have not made reference to many important contributions to the subject, but my objective was not to write a comprehensive review but rather to give an integrated account of one area in which there has been some progress in the past few years, and I apologize to those authors whose work was not mentioned.

## Acknowledgments

I am grateful to Dr. J. N. Fields for a careful reading of the manuscript.

This research was supported by the Division of Basic Energy Sciences, Department of Energy, under contract No. EY-76-C-02-0016.

## *Appendix A: Some Results That Are Useful for Working with Boson Representations*

There are two results that are extremely useful in evaluating correlation functions and commutators when the boson representations of fermion fields are used:

(1) If the commutator $[A, B]$ commutes with $A$ and $B$, then

$$e^{A+B} = e^A e^B e^{-[A,B]/2} \tag{A.1}$$

(2) If $\hat{f}$ is any linear form in boson operators, its thermal average for a free boson distribution is given by

$$\langle \exp(\hat{f}) \rangle = \exp(\tfrac{1}{2}\langle f^2 \rangle) \tag{A.2}$$

The first of these results may be proved by introducing a time-ordering label and writing

$$e^{A+B} = e^A T \exp\left[\int_0^1 dt\, e^{-At} B\, e^{At}\right] \tag{A.3}$$

Then, since $[A, B]$ commutes with $A$, it follows that $e^{-At} B\, e^{At} = B - t[A, B]$ and, substituting this into Equation (A.3), using $[B, [A, B]] = 0$, gives Equation (A.1).

Equation (A.2) may be proved by diagonalizing the free-boson Hamiltonian and deriving the result for each mode separately. The general form is

$$\langle e^{\lambda b^\dagger + \lambda^* b} \rangle \equiv \frac{\mathrm{Tr}(e^{-\beta\omega b^\dagger b}\, e^{\lambda b^\dagger + \lambda^* b})}{\mathrm{Tr}(e^{-\beta\omega b^\dagger b})}$$

$$= (1 - e^{-\beta\omega}) \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} e^{-\beta\omega n} |\lambda|^{2r} \frac{n!}{(r!)^2 (n-r)!} e^{|\lambda|^2/2} \tag{A.4}$$

To obtain the last line, the trace has been evaluated in eigenstates of $b^\dagger b$, Equation (A.1) has been used to separate $\exp(\lambda b^\dagger + \lambda^* b)$ into a product of exponentials and the boson commutation relations have helped to rearrange $\langle (b^\dagger)^r (b)^r \rangle$ as $\langle \hat{n}!/(\hat{n}-r)! \rangle$ where $\hat{n} \equiv b^\dagger b$. If the summation is rearranged so that

$$\sum_{n=0}^{\infty} \sum_{r=0}^{n} a_{nr} \to \sum_{r=0}^{\infty} \sum_{n=r+1}^{\infty} a_{nr}$$

then the summation over $n$ can be carried out to give

$$\langle e^{\lambda b^\dagger + \lambda^* b} \rangle = \sum_{r=0}^{\infty} \frac{1}{r!} |\lambda|^{2r} (e^{\beta\omega} - 1)^{-r} e^{|\lambda|^2/2}$$
$$= \exp[\tfrac{1}{2}|\lambda|^2 (2n+1)] \qquad \text{(A.5)}$$

where $n \equiv (e^{\beta\omega} - 1)^{-1}$ is the Bose occupation factor. Now since

$$\langle (\lambda b^\dagger + \lambda^* b)^2 \rangle = |\lambda|^2 \langle (b^\dagger b + b b^\dagger) \rangle$$
$$= |\lambda|^2 \langle (2 b^\dagger b + 1) \rangle \qquad \text{(A.6)}$$

Equation (A.5) is just

$$\langle e^{\lambda b^\dagger + \lambda^* b} \rangle = \exp[\tfrac{1}{2} \langle (\lambda b^\dagger + \lambda^* b)^2 \rangle] \qquad \text{(A.7)}$$

Then taking the product over all modes gives Equation (A.3).

## *Appendix B: Anticommutation of Different Fermion Fields*

This appendix will consider the method of arranging for different fermion variables to anticommute and its consequences for the transformations used in Section 4. According to the discussion of Section 2, the boson representation ensures that $\psi_{i\sigma}$ and $\psi_{j\sigma'}$ satisfy fermion anticommutation relations for $i = j$ and $\sigma = \sigma'$, but, for $i \neq j$ or $\sigma \neq \sigma'$, it is necessary to include additional phase factors involving number operators as in Equations (59). To specify the convention for operators with spin, the operators will be labeled $\psi_m(x)$ with $m = 1, 2, 3, 4$ denoting $\psi_{1\uparrow}$, $\psi_{2\uparrow}$, $\psi_{1\downarrow}$, $\psi_{2\downarrow}$ in that order. Then the phase factor for $\psi_m$ is taken to be $\exp[(i\pi/2)(\sum_{n=1}^{4} \varepsilon_{mn} \hat{N}_n)]$, where $\varepsilon_{mn}$ is an antisymmetric matrix with $\varepsilon_{mn} = 1$ for $n > m$. Then, commuting $\psi_p \psi_q$ for $p \neq q$ gives a factor $\exp(i\pi\varepsilon_{pq}) = 1$ as required. Inserting this convention into the backward scattering operator (112), the phase factor is $\exp[(i\pi/2)(N_{1\downarrow} + N_{2\downarrow} - N_{1\uparrow} - N_{2\uparrow})]$ whereas, for the umklapp scattering operator (111), it is $\exp[(i\pi/2)(N_{1\uparrow} + N_{2\uparrow} + N_{1\downarrow} + N_{2\downarrow})]$. Since both of these exponents commute with $H$ and all answers depend only on the modulus of



the coupling constant, the phase factors have been omitted from Equations (112) and (113) and, by a similar argument, from Equations (115) and (116).

## *Appendix C: Charge-Density Wave Gap and CDW Correlations*

The arguments of Sections 4 and 5, if followed literally, would lead to a CDW exponent of $\theta_{sR} - 2$ for $\bar{W}_\parallel > -|\bar{W}_\perp|$ and $\bar{U}_\parallel \geq |\bar{U}_\perp|$ when there is a half-filled band, and this would allow the formation of long-range CDW correlations despite the presence of a gap in the charge-density spectrum, in qualitative disagreement with the conclusions for the strong coupling limit.

This question may be examined in another way for the Hubbard model, which has $U_\parallel = U_\perp = W_\parallel = W_\perp = sU$ (see Table 1) and falls in the region of consideration for $U > 0$. Since $\bar{U}_R = 0$ in this case, the straightforward argument gives a CDW exponent of $-1$. On the other hand, for a half-filled band, the transformation

$$a_{i\uparrow} \to a_{i\uparrow}$$
$$a_{i\downarrow} \to a_{i\downarrow}^\dagger \qquad \text{(C.1)}$$

has the effect of reversing the sign of $U$ in Equations (4) and (5) without changing the number of particles. The same transformation effects the following interchanges of correlation functions:

$$\text{SP} \leftrightarrow \text{transverse SDW}$$
$$\text{CDW} \leftrightarrow \text{longitudinal SDW} \qquad \text{(C.2)}$$

Table 3, as written, satisfies this symmetry since $\theta_{cR} = \theta_{sR} = 1$ for the Hubbard model, but it would not have done so if the original naive CDW exponent for $U > 0$ had been used.

Physically it is reasonable that, if an SDW gap prevents an SDW power law for $U < 0$, then a CDW gap should have the same effect on the CDW correlations when $U > 0$, and this conclusion should not be special to the Hubbard model. In the strong coupling limit, spin symmetry is satisfied by having long-range correlations in the up-spins and down-spins separately, but they are exactly out of phase so that one compensates the effect that the other produces on the charge density. For weaker coupling, it is easier in general for CDW and SDW order to coexist, but there is little doubt that the argument for the effect of a gap is less fragile for the SDW