CONF-9107158--1

DE91 016759

Proceedings of
Symposium on Electron Correlation
Trieste, Italy
July 22–August 3, 1991

ELECTRON ENERGIES IN METALS

G. D. Mahan

Department of Physics and Astronomy,
University of Tennessee, Knoxville, TN

and

SOLID STATE DIVISION
OAK RIDGE NATIONAL LABORATORY
Managed by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
Under
Contract No. DE-AC05-84OR21400
With the
U. S. DEPARTMENT OF ENERGY
OAK RIDGE, TENNESSEE

July 1991

Electron Energies in Metals

G.D. Mahan
Department of Physics and Astronomy
The University of Tennessee
Knoxville, TN 37996-1200, and

Solid State Division, Oak Ridge National Laboratory
P.O. Box 2008, Oak Ridge, TN 37831-6030

July 10, 1991

1 Introduction

The modern theory of metals began with the free electron theory of Sommerfeld[1]. He proposed that the valence electrons behave as a noninteracting gas of unpolarized electrons. This predicts that the Fermi degeneracy is given by $E_F = \hbar^2(3\pi^2 n)^{2/3}/2m$, where the electron density is n. Early soft x-ray emission experiments by Skinner[2] seemed to confirm this picture. Also see discussion by Seitz[3]. The band widths he observed were in good agreement with the predictions of free electron theory. We call this period the ancient era of electron-electron interactions.

The middle ages of electron-electron interactions began about 1960. Then calculations were started on the self energies of individual electrons, and on the ground state energy of the systems of interacting electrons and ions. Gradually, an understanding was achived regarding electron correlation. Hubbard[4] introduced the concept of the local field factor $G(q)$, which was calculated by a number of techniques. The best early work was certainly by Singwi and his colleagues[5-8]. They developed a self-consistent set of equations which determined the local field expression, and calculated it as a function of wave

vector q and electron density. Later work by Ceperley[9-10] and others largely confirmed the correctness of this pioneering work of Singwi's group.

Electron self-energy calculations were done by Hedin[11] and by Rice[12]. Hedin used RPA, while Rice used a Hubbard suggestion for $G(q)$. The two approaches agreed for most of the quantities. For example, they found that electron-electron interactions made only a small contribution to the effective mass of the electrons in the occupied bands. The interactions are large, but are not particularly k-dependent.

Many experiments confirmed this quasi-particle picture. The Fermi surface of simple metals was determined by various cyclotron resonance and optical conductivity experiments. For the simple metals such as Na, Mg, Al, K, etc., they found that the effective masses of the electrons were near to the free electron masses. The differences were ascribed to the electron-phonon interaction. This latter contribution only affects electrons very near in energy to the Fermi surface, and had no effect on the majority of band electrons. Thus these detailed calculations tended to support the free electron hypothesis of Sommerfeld.

The modern era of electron-electron interactions began a decade ago. Plummer's group[13-17] initiated a program of using angular resolved photoemission to examine the band structure of the simple metals. Beginning with aluminum, and carrying on to sodium and potassium, they always found that the occupied energy bands were much narrower than expected. For example, the compressed energy bands for metallic potassium suggest a band effective mass of[17] $m* = 1.33m_e$. This should be commpared to the band mass found from optical conductivity [18] $m*/m_e = 1.01 \pm 0.01$. The discrepancy between these results is startling. It was this great difference which started my group doing calculations. Our program was two-fold. On one hand, we reanalyzed the experimental data, in order to see if Plummer's result was an experimental artifact. On the other hand, we completely redid the electron-electron self-energy calculations for simple metals, using the most modern choices of local-field corrections and vertex corrections. Our results will be reported in these lectures. They can be summarized as following: Our calculations give the same effective masses as the older calculations, so the theory is relatively unchanged; Our analysis of the experiments suggests that the recent measurements of band narrowing are an experimental artifact.

2 Experiments on Band Narrowing

2.1 Photoemission

In 1985 Jensen and Plummer[13] announced their photoemission experiments had shown the existence of charge density waves in metallic sodium. That spectacular assertion was the reason we started analyzing their data. Since then, as a result of our calculations[19–22], they have retracted the claim about charge density waves in sodium.

If they had claimed to see charge-density waves (CDW) in potassium, we would not have bothered. That metal has a number of peculiar properties, especially its photoemission[17]. The existence of CDW in potassium is eternally controversial. However there is no other real evidence for CDW behavior in metallic sodium.

However, their work announced another experimental discovery. They found the electron energy band width of sodium was 10% narrower than thought previously. This latter discovery has received support from other experiments. In this section we will review these experimental measurements. Our own view is quite equivocal. We have reanalyzed each different experiment, and find in each case that the evidence for unexpected band narrowing is doubtful. On the other hand, the experimentalists are each quite capable, and we not wish to assert that our interpretation is necessarily better than theirs. As the tide of opinion rises for band narrowing, we seem to be the only voice of complaint. Perhaps everyone else is right.

Fig.1 shows some of the experimental results of Jensen and Plummer[13]. They measured the energies of electrons leaving normal to the (011) surface. Fig. 1(a) shows the energy distribution curves. For each photon energy, they identified the peak height as marking the interband transition. They also know the final wave vector of the electron, since it is normal to the surface. Then they deduced the initial energy and wave vector of the electron, which permitted them to map the energy bands as shown in Fig. 1(b). Calling the zero of energy as the Fermi surface, the bottom of their measured band was at $E = -2.5$eV instead of the free electron value of $E = -3.14$eV. Until that time, the expected band width of sodium and other metals is shown in Table 1. The top line is the prediction in electron volts of free electron theory $E_{F0} = \hbar^2(3\pi^2 n)^{2/3}/2m$. We use the density n at room temperature,

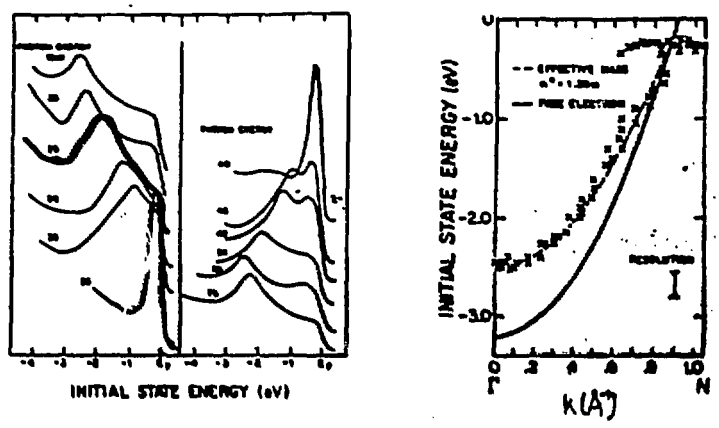

Figure 1: photoemission results of Jensen and Plummer. (a) energy distribution curves for electrons leaving normal to the surface. (b) measured band narrowing (points) compared to free electron theory (solid line).

Metal	Na	Mg	Al
Free electron	3.14	7.11	11.66
Band structure	0.	-0.2	-0.7
Electron-electron	-0.35	-0.27	-0.04
Total	2.79	6.64	11.0

Table 1: Band width energies in eV. Top line are predictions of free electron theory $\hbar^2(3\pi^2 n)^{2/3}/2m$ where n is the density of conduction electrons. The second line is the change caused by band theory: one electron energies in the periodic potential of the ions. The third column are our electron-electron calculations.

	Na	K
optical	1.01±0.01[a]	1.01±0.01[b]
conductivity	1.07±0.01[c]	1.08±0.01[c]
cyclotron resonance	1.24±0.02[d]	1.21±0.02[d]
photoemission	1.23[e]	1.33[f]

a. JN Hodgson, J. Phys. Chem. Solids **24**, 1213 (1963)
b. US Whang, ET Arakawa, TA Callcott, Phys. Rev. B **6**, 2109 (1972)
c. J Monin and GA Boutry, Phys. Rev. B **9**, 1309 (1974)
d. CC Grimes and AF Kip, Phys. Rev. **132**, 1991 (1963)
e. IW Lyo and EW Plummer, Phys. Rev. Lett. **60**, 1558 (1988)
f. BS Itchkawitz et al., Phys. Rev. B **41**, 8075 (1990).

Table 2: Effective masses $m*/m_e$ in Na and K by various measurements.

since most measurements are done there. The second line are the change in the band widths found in energy band calculations. This change increase with the ion valence, and is large in aluminum. The third line lists our[22] electron-electron corrections, which are similar to those predicted long ago by Hedin using RPA. For all three metals, the band theory plus the electron-electron interactions make a sizeable correction to the band width. This table is what one would have constructed in 1985, before the current round of experiments and theory. Jensen and Plummer's value of 2.5 eV for sodium is more than 10% smaller than the expected result. Plummer's group also reported narrowing for aluminum and potassium.

Table 2 shows the effective masses measured in sodium and potassium by various measurements. The optical conductivity determines the band masses. Several measurements give slightly different values. The cyclotron resonance experiment measures the band masses enhanced by the electron-phonon contribution $(1+\lambda)$. For Na and K the values of $\lambda \sim 0.15$ which explains why the cyclotron masses are higher than the optical masses. The recent photoemission experiments measure band masses, which do not contain the factor of $(1+\lambda)$. There is no way to reconcile the large values from photoemission with the results of the other measurements.

Shung and I did some one-electron calculations for the photoemission of sodium (011). We used my 1970 theory[23] which included:

- surface potentials $V(z)$ for the electrons taken from Lang and Kohn,
- numerically generated eigenfunctions and matrix elements,
- interference between the surface and volume effect,
- and the finite mean free path of the electron in the final state.

Later we also included the mean free path of the electron in the inital state. It is important to realize that at the photon energies of the experiment, the mean free path of the electrons leaving the surface is about 5-6 Å. This short value ensures that one measures only electrons from the first one or two atomic layers of the surface. One might expect that surface effects are important in interpreting the results.

Fig. 2 shows an energy distribution curve for a photon energy of $\hbar\omega =$ 24eV. The interband transitions peak at the point labeled A. This peak is broad because of the short mean-free path (MFP). The surface effect gives a small contribution. More interesting is the interference between the surface and volume effect, which is the dot-dashed line. The total is the heavy black line. We then broadened it with the experimental resolution, and produced the dashed line. It has a peak at the point labeled B. There is a systematic shift of 0.2-0.3 eV between the peaks A and B. Point A is where the interband transition occurs, while B is where the peak has its maximum. They differ because of the interference between the surface and volume processes in photoemission. Interpreting the peak position as the interband point would make the bands appear narrower than they are. This explains the apparent narrowing first observed by Jensen and Plummer. In using peak heights to identify the position of the interband transition, they had a systematic error which make the bands appear narrower. Our calculations show that this difference persists over a wide range of photon energies.

The calculated energy distribution curves are quite square in shape, which agrees with the measurements shown in Fig. 1a. We also explain numerous other features of the experiments, such as the decrease in intensity near band gaps, and the dependence upon the polar angle when the measuring cup is swung through a polar angle.

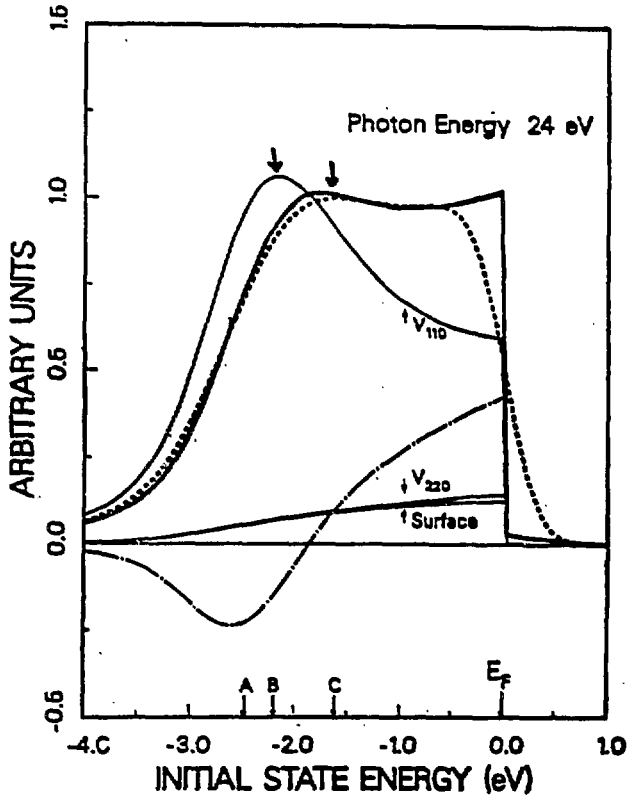

Figure 2: Energy distribution curves for sodium (011) at a photon energy of $\hbar\omega = 24$ eV calculated by Shung and Mahan[19–20]. The thin solid line shows the interband contribution alone. It peak denotes the interband transition. The heavy solid line combines the surface and volume effects. The peak position is 0.2-0.3 eV higher, because of surface effects. Identifying this as the interband point would lead to apparent narrowing, which is fact only interference.

Lyo and Plummer[16] did further experiments on simple metals. Instead of measuring the electrons leaving normal to the surface, they mapped the energy bands by finding the critical polar angle at which the emission disappears. They claim this is a more accurate way of mapping the bands. Their value for the band width for sodium is 2.65 eV. This is higher than the 2.5 result of Jensen and Plummer, and is getting close to the value of 2.79 expected from traditional theory. No one has calculated the role of surface interference at the large polar angles of their experiment, so it is unknown whether their anular measurements are affected in the same way, by interference, as are the measurements normal to the surface.

In this letter, Lyo and Plummer also claimed that excessive band narrowing is expected from electron-electron interactions. However, their calculations had two major approximations: the plasmon pole approximation, and the neglect of vertex corrections. As discussed below, our group redid these calculations without either of these approximations, and found that the results changed appreciably. In fact, their calculations were simply misleading. Our calculations are discussed below.

The energy of each electron is affected by electron-electron interactions with the other electrons. The largest term is called the screened exchange interaction $\Sigma_{sx}(p, E)$. The technique for evaluating this quantity is discussed in later sections. Here we wish to show calculations for the simple metals, and discuss their relevence to photoemission measurements. Fig.3 shows the screened exchange energy for simple metals for $E = \varepsilon_p$. Now just examine the real part of the self energy for sodium. At $p = 0$, the band minimum, it is about -5.5 eV, which is twice the Fermi energy! That is large. There is only a small dependence upon p between $p = 0$ and k_F. To a good approximation, the band is shifted rigidly downward by this amount, at least for the electron states in the Fermi sea. However, the excited electron in photoemission has a screened exchange energy which gets smaller in magnitude at large wave vector. The dip in the self energy comes from plasmon emission, and occurs one plasmon energy above the Fermi surface. Energy conservation in photoemission is

$$\hbar\omega + \varepsilon_k + \Re\Sigma(k) = \varepsilon_{\vec{k}+\vec{G}} + \Re\Sigma(\vec{k}+\vec{G}), \quad (1)$$

$$\varepsilon_k = \frac{\hbar^2 k^2}{2m}, \quad (2)$$

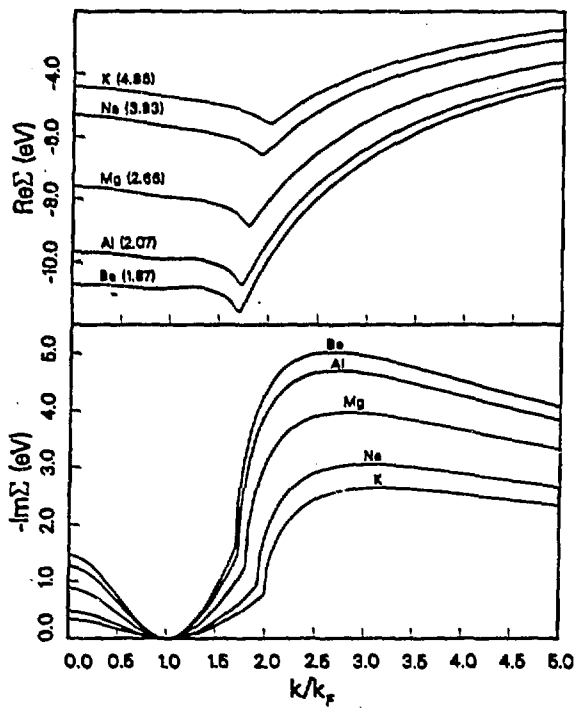

Figure 3: Real and imaginary part of the screened exchange energy evaluated at energies $E = \xi_p$ calculated by Shung, Sernelius, and Mahan[21].

where $\vec{G}$ is a reciprocal lattice vector. In order to visualize the self energy, examine Fig.4. There we show the energy bands when electrons are constrained to move in the (011) direction. This is the band structure measured by Plummer's group when they capture only electrons exiting normal to the (011) surface. The solid lines are the free electron bands, while the dashed line are the bands including the screened exchange energy. We have set the self energy to be zero at the Fermi surface. One does that in practice, since the chemical potential is renormalized by the self energy $\Re\Sigma(k_F)$. At point A, the band minimum is slightly higher. This is the shift of -0.35 eV given in Table 1. At the point B the self energy shift, relative to the Fermi surface, is almost zero. A transition from A to B, by photon absorption, requires a smaller energy than predicted by free electron theory because of the self energy difference between these two points. Plummer's group measures the band minimum by locating this point. They neglect these self energy shifts. This neglect makes the band minimum appear to be shifted upwards. Our calculated difference between the self-energies at A and B is only about 0.1 eV, which is a small correction. However, between the points A and C the self-energy difference is about 2.4 eV. Thus, the self-energy effects would make the band minimum to appear deeper by 2.4 eV based on interpreting the transition from A to C with a free electron model. This is not a small effect: the shift is nearly equal to the Fermi energy itself!

Plummer's group finds the same band narrowing from the A to B as from A to C. That is, they find no evidence for a self-energy shift. We find the absence of these self-energy effects to be remarkable.

Shung, Sernelius, and Mahan (SSM) showed that these self-energy effects did affect the two transitions marked D and E. Since the Fermi wave vector k_F does not go to the zone edge, their should be a gap in the photoemission spectra between photon energies D and E. This gap is found experimentally. The photoemission does not go to zero in this region, but is very small. The nonzero value is due to the finite MFP of the electron. SSM showed that the free electron theory predicts the wrong position for this spectral minimum, compared with experiment. However, including the 1.4 shift caused by difference in the screened exchange energy brings the spectral gap into good agreement with the experiment.

In conclusion, the experiments show the screened exchange energy exists

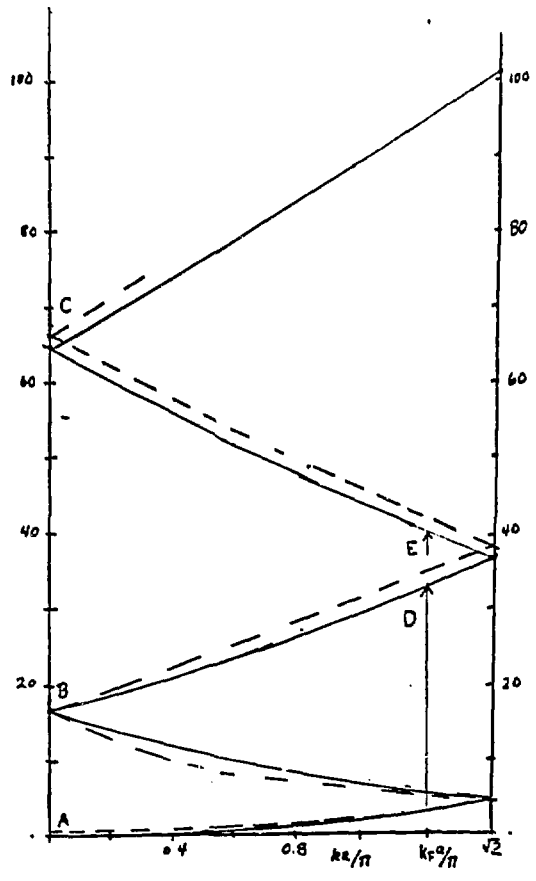

Figure 4: Energy bands in Na projected in the (011) direction. The solid lines are free electron bands, while the dashed line are the bands including the screened exchange energy.

on the right side of the graph in Fig.4, but not on the left side. Of course, this is nonsense.

2.2 Soft X-ray Emission

Soft x-ray emission spectroscopy is the historical way of measuring the energy widths of occupied electron bands in solids. The early work of Skinner[2-3] confirmed that the band widths were similar to the predictions of free electron theory. Recent emission spectrometers have been built with much greater sensitivity than before, which permits many new kinds of measurements. However, the emission spectra of simple metals have not actually changed much from Skinner's results.

What has changed is the interpretation of the experiments. For example, in 1974 Neddermeyer[24] measured the emission spectra of aluminum and reported that the band width was 11.67±0.14 eV. Recently Livins and Schnatterly[25] repeated the same experiment, and reported that the aluminum band width was 10.6 eV. It is important to realize that the spectra they measured were identical. What differed was the interpretation. Neddermeyer's result is the measured width. Livins and Schnatterly attempted to apply many-body corrections to the interpretation of the data. We agree with them that many-body effects cause widening of the measured spectra, and that a correction should be applied which will produce a smaller width than the raw data. We congratulate them for being the first to attempt this correction, but disagree with their method of reducing the data. The many-body corrections they analyze come from electron-electron interactions. We have reanalyzed this data, and deduce an experimental width from emission spectroscopy of 11.0 eV for aluminum. Our number is in good agreement with the theoretical result in Table 1. The band are narrower than the free electron value, but the reduction is from energy band theory, rather than from electron-electron self-energies. However, Plummer's photoemission measurements[14] have been interpreted as showing band narrowing in aluminum. They also like the number of 10.6 eV. However, there is still the issue of how much the photoemission results are influenced by surface effects.

Fig.5 shows the L-emission spectra of aluminum. The sharp rise on the right is at the Fermi surface. On the left, the spectral intensity smoothly goes to zero. Where, in this tail, is the band minimum? Obviously, to find the

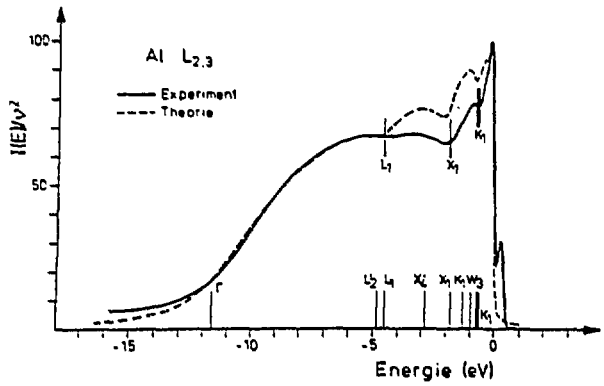

Figure 5: L emission spectra of metallic aluminum, from Neddermeyer[24].

width one has to find two points: the band bottom and the Fermi surface. The Fermi surface is sharply defined in the spectra. However, finding the band minimum is a problem because the spectra goes to zero smoothly. The data can only be interpreted if one understands this spectral region.

Two processes contribute to the spectral shape in the band tail. One is the lifetime of the electron at the band minimum. The most important contribution to the lifetime is electron-electron scattering, which gives a value of about $\hbar/\tau = 1.3eV$. This contribution is obviously important. Another contribution to the broadening is caused by the measurement process. During the emission event, an electron fills a core hole, and changes the external valence of the ion. The other electrons in the conduction band respond by creating numerous electron-hole pairs. This process is part of the *orthogonality catastrophe* first discussed by Anderson[26]. It has a big effect on the spectral shape of emission. It was this process which was included by Livins and Schnatterly.

X-ray edge singularities occur at the Fermi edge in x-ray absorption and emission spectroscopy. The phenomena is described by a simple Hamiltonian known as the MND problem[26]. Several exact solutions have been derived for this model[27–29]. The most recent solutions are also easy to calculate, so it can be said that the problem now has a simple, exact solution. This solution applies to the entire spectral region. It also includes the region of the band

tail. It is in this region where the band width controversies occur in soft x-ray emission. The exact solutions to the MND problem are discussed elsewhere. Here we shall only mention one feature of the result.

Assume that one has a good computer code which calculates energy bands, optical matrix elements, and the entire emission spectra $A_0(\omega)$ according to band theory. Now one must add the MND effects to the spectral predictions. This process is rather easy, since there is a convolution function $K(\omega)$ such that the predicted spectra is[30]

$$A(\omega) = \int d\omega' K(\omega - \omega') A_0(\omega'). \tag{3}$$

The function $K(\omega)$ conserves area, and its integral over all frequency gives zero. In this way one can predict experimental spectra, including many-body theory. However, the experimentlist needs something else. They measures $A(\omega)$, and wants to find the band theory prediction $A_0(\omega)$. Thus they need the inverse operator

$$A_0(\omega) = \int d\omega' K^{-1}(\omega - \omega') A(\omega'). \tag{4}$$

However, the operator $K(\omega)$ is quite singular. We have been unable to construct its inverse. It needs to be done, and remains a challenge for theorists.

2.3 X-Ray Absorption

Another experiment which claims to show band narrowing is x-ray absorption. Both Citrin et al.[31], and Callcott et al.[32], claim to see evidence for narrowing in metallic sodium. Here the key is the critical point in the density of states at the zone edge. This critical point puts a well-defined signature in the x-ray absorption. A careful measurement of the frequency difference from this point, and the Fermi edge, gives a measuremnt of the band narrowing over this region. Because of spin-orbit splitting of the L-core state, the absorption spectra should actually see two critical points, depending upon whether the initial state is $p_{3/2}$ or $p_{1/2}$. Unfortunately, only one critical point appears in the data, as shown in Fig.6. Whether band narrowing occurs, or does not occur, depends upon which spin-orbit component you assign the critical point. As with everything in this field, nothing is certain.

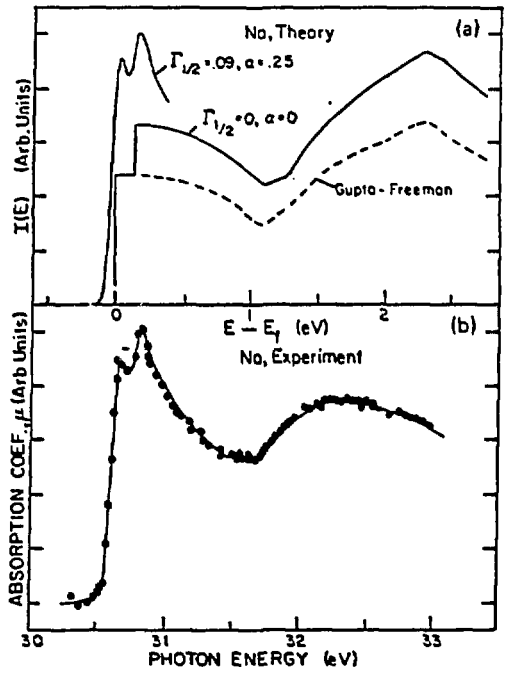

Figure 6: X-ray absorption of metallic sodium. Top shows theoretical prediction, incuding the two components split by the spin–orbit interaction. Bottom shows experimental data from Callcott et al. [32].

3 Dielectric Functions

The conduction electrons in a metal contribute to the dielectric screening. In crystalline solids, the screening is provided by a dielectric function $\varepsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}', \omega)$, where $\vec{G}$ and $\vec{G}'$ are reciprocal lattice vectors[33]. The double argument for wave vector arises because at an atomic level, the dipole moment at point $\vec{r}$ from a field applied to point $\vec{r}'$ is not a function of $(\vec{r}-\vec{r}')$. Atomic polarization is more nonlocal than that, and so is the polarization in solids at an atomic level.

The homogeneous electron gas is a mathematical model for a simple metal. Here the ion charges are spread uniformly throughout the solid as a uniform positive background. There are no ions nor atoms. One aspect of the model is that the dielectric response is now a function of $(\vec{r}-\vec{r}')$. After Fourier transforming, the dielectric function can be written as $\varepsilon(q, i\omega)$. The formal definition for the homogeneous electron gas is

$$\frac{1}{\varepsilon(q,i\omega)} = 1 - \frac{v_q}{\Omega}\int_0^\beta d\tau e^{i\omega\tau} < T_\tau \rho(q,\tau)\rho^\dagger(q,0) >, \tag{5}$$

$$\rho(q) = \sum_{\vec{k},\sigma} C^\dagger_{\vec{k}+\vec{q},\sigma} C_{\vec{k},\sigma}, \tag{6}$$

where $v_q = 4\pi e^2/q^2$ in three dimensions, and Ω is the volume. We are using the nonzero temperature formalism, in which the frequencies are $i\omega_n$, and we suppress the subscript n. It is important to note that the definition of the dielectric function is, in fact, the definition of the inverse dielectric function. One can show that the retarded form $(1/\varepsilon(q,\omega))$ is a causal function, which means that it has no poles in the upper half plane of frequency space. The dielectric function $\varepsilon(q,\omega)$ is not causal, and it could have poles in the upper half plane. Since the inverse dielectric function is causal, we can express it in a Lehmann representation, in terms of its spectral function $B(q,\omega)$,

$$B(q,\omega) = -2\Im\left(\frac{1}{\varepsilon(q,\omega)}\right), \tag{7}$$

$$\frac{1}{\varepsilon(q,Z)} = 1 + \int_\infty^\infty \frac{d\omega'}{2\pi}\frac{B(q,\omega')}{Z-\omega'}. \tag{8}$$

Here the frequency z can be either the complex one of $i\omega_n$ or else the real frequency $\omega + i\delta$, where δ is infinitesimal. One interesting case is at zero

frequency $Z = 0$. Since $B(q, \omega)$ is antisymmetric in frequency, one can write at zero temperature

$$\frac{1}{\epsilon(q,0)} = 1 - \int_0^\infty \frac{d\omega'}{\pi} \frac{B(q,\omega')}{\omega'}. \tag{9}$$

The integral is strictly positive. So at zero frequency one can prove that

$$\frac{1}{\epsilon} \leq 1. \tag{10}$$

Note that being ≤ 1 permits negative values. Thus, we conclude that $\epsilon(q)$ must be greater than one, or less than zero, but it can never have a value between zero and one. This constraint is well-known, and is a result of causality[33]. Usually we do not have an exact expression for the dielectric function, and instead must use approximate expressions. One must alway ensure that one's approximation does not violate any of these obvious constraints. There are also numerous sum rules on the dielectric function.

For the homogeneous electron gas, the first dielectric function which contained frequency response was constructed by Lindhard. Today we call it the *random phase approximation*, or RPA[26,34]. Students always ask which phases are random. I never understood the name either–read Pine's book[34] for an explanation. The paramagnetic susceptibility of the electron gas is given in terms of a polarization function $P(q, i\omega)$, which is

$$\epsilon_{RPA}(q, i\omega) = 1 - \chi(q, i\omega), \tag{11}$$

$$\chi(q, i\omega) = v_q P(q, i\omega), \tag{12}$$

$$P(q, i\omega) = 2 \int \frac{d^3p}{(2\pi)^3} \frac{n_{\vec{p}} - n_{\vec{p}+\vec{q}}}{i\omega + \varepsilon_{\vec{p}} - \varepsilon_{\vec{p}+\vec{q}}}, \tag{13}$$

where $n_{\vec{p}}$ is the occupation number of the electrons. For quadratic dispersion ($\varepsilon_{\vec{p}} = p^2/2m$) the d^3p integral can be done analytically at zero temperature. The results are listed in thick books[26]. The same expression for $P(\vec{q}, i\omega)$ is valid for tight-binding models of solids. There the band disperison is not parabolic, and the d^3p integrals can not be evaluated analytically. Here we shall keep the results simple, and only evaluate the case for parabolic dispersion.

Hubbard	$\frac{0.5x^2}{1+x^2}$
Geldart–Vosko	$\frac{0.5x^2}{2+x^2}$
Vashishta–Singwi	$A(1-e^{-Bx^2})$

Table 3: Expressions for $G(q)$, where $x = q/k_F$.

RPA is a good dielectric function. However, it is an approximation, and some quantities calculated with it come out poorly, such as the pair distribution function $g(r)$. Hubbard suggested a variation on RPA, which is a major improvement and has become universally adopted. Hubbard[4] suggested adding a local field factor $G(q)$ to the dielectric function,

$$\varepsilon(q,i\omega) = 1-\chi(q,i\omega)\Gamma(q,i\omega), \tag{14}$$

$$\Gamma(q,i\omega) = \frac{1}{1+G(q)\chi(q,i\omega)}. \tag{15}$$

Now it is known that one can rigorously write the dielectric fashion in this way, with the proviso that the local field factor $G(q,i\omega)$ also depends upon frequency. A number of exact results[35–37] have been derived for this function, and some of them are:

$$G(q) = \frac{1}{2N}\sum_{q'\sigma\sigma'}\left[\frac{(\vec{q}\cdot\vec{q}')^2}{q^2q'^2}-1\right][S_{\sigma\sigma'}(q')-\delta_{\sigma\sigma'}], \tag{16}$$

$$S_{\sigma\sigma'}(q) = \frac{2}{N}<|\rho^\dagger_\sigma(q)\rho_{\sigma'}(q)|>, \tag{17}$$

$$\lim_{q\to\infty} G(q,\omega) = 1-g(0). \tag{18}$$

$$\tag{19}$$

The expression for $G(q)$ in terms of the *static structure factor* $S(q)$ was first given by Singwi et al[5–8]. For our calculations we use static approximations to $G(q)$. Several standard approximations are listed in table 3. All approximations have the feature that $G(q) \sim q^2$ at small values of q, and $G(q)$ becomes a constant at large values of q. In the Vashishta–Singwi[7] expression, the dimensionless parameters A and B are functions of electron density. Usually for our calculations we employ either the Geldart–Vosko[38] form, or else the Vashishta–Singwi form. For the quantities we calculate, the results are nearly identical when using these two forms.

4 Vertex Corrections

The self-energy of an electron from electron-electron interactions with other electrons can be written, for the homogeneous electron gas, as the formula:

$$\Sigma(p) = \int dq W(q)\Gamma(p, p+q)\mathcal{G}(p+q). \tag{20}$$

A four vector notation is used, where at nonzero temperature the symbol $q = (\vec{q}, iq_n)$, and the integral dq means a three dimensional integral over the wave vector, plus a summation over the descrete frequency variable iq_n. The symbol $W(q) = v_q/\varepsilon(q)$ is the screened interaction. The electron's Green's function is $\mathcal{G}(p+q)$. Do not confuse the Green's function with the local field factor. The vertex correction is $\Gamma(p, p+q)$.

The message of this section is rather simple: For dielectric functions of the Hubbard type, the correct vertex function depends only upon (q) and is

$$\Gamma(q) = \frac{1}{1+\chi(q)G(q)}, \tag{21}$$

$$\tilde{W}(q) = W(q)\Gamma(q) = \frac{v_q}{1-\chi(q)(1-G(q))}. \tag{22}$$

The combination of $\Gamma(q)/\varepsilon(q) \equiv 1/\tilde{\varepsilon}(q)$ produces an effective dielectric function which is $\tilde{\varepsilon}(q) = 1-\chi(q)(1-G)$. It turns out this expression is not very sensitive to the form of the local field factor $G(q)$.

Another important point is that the correct electron's Green's function is the noninteracting one $\mathcal{G}_0(p+q)$. That is not obvious. In fact, it is more-or-less a definition. The form for the local field factor $G(q)$ is derived assuming that all Green's functions are noninteracting. Thus the local field factor contains some correction terms due to the electron self-energy. Rather than include them in the Green's function, one has instead included them in the local-field factor. This procedure is sensible because there is a great deal of cancellation between vertex corrections and self-energies. Put both factors into the local-field factor, where they cancel and yield a small correction.

Although the above results are simple, they are misunderstood to the point of being controversial. Thus we shall spend a little time here discussing them.

The confusing point is on the nature and number of vertex corrections. The answer is that the vertex correction only occurs as a single factor in

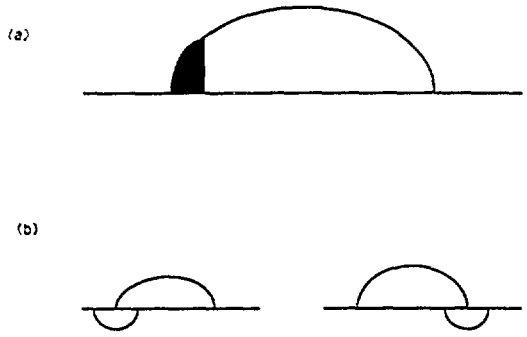

Figure 7: Feynman diagram showing electron self energy. Horizontal line is electron Green's function. Loops are screened electron-electron interaction. Shaded area is vertex correction

the expression (20) for the self-energy. This result has been understood, and unchanging, from the early days of field theory[39] . However, people still get it wrong.

Figure 7a shows a Feynman diagram of the self-energy in (20). The horizonal line represents the electron Green's function. The upper loop represents the screened interaction $W(q)$. The shaded area represents the vertex correction $\Gamma(q)$. The question is whether the vertex correction should dress one vertex, or two. That is, should there be shading, and another factor of $\Gamma(q)$, at the the other place where the loop meets the horizontal line, on the right. The answer is 'no'.

Figure 7b shows two Feynman diagrams. The first is obviously a vertex correction to the left vertex. The other is a vertex correction to the right vertex. In fact, these two diagrams are identical. In perturbation theory, the diagram occurs only once, not twice. So already, in the first order of perturbation theory, the message is that only one vertex is being renormalized.

A test particle[40] is defined as a classical particle of charge $Q << e$, where e is the charge on an electron. Since the particle is classical, it has no exchange interactions with the electrons. The dielectric function $\varepsilon(q)$ is defined rigorously as the screening, in the electron gas, between two test

particles,

$$V_{tt}(q) = \frac{4\pi Q^2}{q^2 \varepsilon(q)}. \tag{23}$$

How is this interaction altered when the test particles are replaced by electrons? Proceeding one step at a time, first consider the interaction V_{te} between a test particle and an electron. Now there are exchange effects between the electron, and those electrons providing the screening. This changes the effective interaction, and this change is a 'vertex correction',

$$V_{te} = \Gamma(q) V_{tt}(q). \tag{24}$$

The vertex correction is the renormalization of the interaction caused by exchange effects between the electron, and those electrons providing the screening. The next step in the argument is to consider the interaction between two electrons. The two electrons have exchange interactions with each other, which makes the effective interaction spin-dependent. However, the argument proceeds by renormalizing the interaction for each of the two electrons. For electrons with parallel spin, this reasoning gives $V_{ee} = \Gamma^2 V_{tt}$. This result is incorrect, as we mentioned above. In evaluating the self-energy of the electron, the correct expression is V_{te}. This does not seem sensible intuitively, but it is correct.

So far we have given the argument about why the vertex correction only occurs once. Next, why do we use the particular form in (20)? The best argument for this was given by Ting, Lee, and Quinn[41]. They used the fact that the functional derivative of the ground state energy, with respect to the particle occupation number, yields the single particle energy

$$\frac{\delta E_G}{\delta n_p} = E_p = \varepsilon_p + \Sigma(p), \tag{25}$$

$$E_G = E_{G0} - \frac{\Omega}{2\beta} \int_0^1 \frac{d\eta}{\eta} \int \frac{d^3q}{(2\pi)^3} \sum_{iq} \left[\frac{1}{\varepsilon(q, iq)} - 1 \right], \tag{26}$$

$$\varepsilon(q, i\omega) = 1 - \frac{\eta \chi(q)}{1 + \eta G(q) \chi(q)}. \tag{27}$$

The dependence upon n_p occurs in the electron polarization $P(q)$. Evaluating the functional derivative, and then doing the coupling constant integration

$d\eta$, yields the expression (20) with $ip = \xi_p$. One also gets this expression by summing the diagrams which constitute the vertex correction. The same diagrams that renormalize $P(q)$ also renormalize the electron self-energy.

As an aside, we also comment on how vertex corrections normalize the self-energy of the electron from the electron-phonon interaction. Here the self-energy can be written as

$$\Sigma(p) = \int dq \tilde{U}(q)\mathcal{D}(q)\mathcal{G}(p+q), \tag{28}$$

$$\tilde{U}(q) = \left|\frac{M(q)}{\varepsilon(q)}\right|^2 \Gamma^2(q) = \left|\frac{M(q)}{\tilde{\varepsilon}(q)}\right|^2. \tag{29}$$

The effective interaction $\tilde{U}(q)$ contains the square of the electron-phonon matrix element, divided by the dielectric constant. Again one uses the renormalized screening $\tilde{\varepsilon}(q) = 1 - \chi(1-G)$. The vertex correction does occur twice in this case. The reason it occurs twice here, and only once in electron-electron interactions, is that now the loop part of the diagram is different–it is a phonon whose Green's function is $\mathcal{D}(q)$. So one can renormalize each end. When everything is electron-electron interactions, one only renormalizes one end. Check it out by counting diagrams.

5 Numerical Results

The electron self-energy has been evaluated numerically, and here we present some results. In a recent paper, the authors commented[42] that "No special difficulty is present in performing such a calculation". They are wrong. As evidence, I offer the observation that those authors published inaccurate numbers. Three different members of my group have written separate computer codes, and we agree to order ± 0.1mRy. My own code is the least accurate of the three.

Equation (20) contains four integrals. The angular integrals $d\phi d\theta$ in d^3p can be done analytically. The dp integral is done numerically. At zero temperature the summation over ip_n is changed to a continuous integral over real and imaginary frequency, as described elsewhere. In the end, one has to evaluate numerically a double integral over wave vector and frequency.

The numerical problems are two fold. First, neither integral converges rapidly at large values. One must grind out a lot of points. Secondly, the

$r_s =$	2	3	4	5
RPA				
$p = 0$	0.7364	0.4993	0.3854	0.3180
$p = k_F$	0.7499	0.5270	0.4123	0.3419
VS				
$p = 0$	0.7059	0.4619	0.3452	0.2766
$p = k_F$	0.7087	0.4859	0.3714	0.3011

Table 4: Electron self-energies $-\Re\Sigma(p, \epsilon_p)$ in rydbergs. Values given at the bottom of the band and at the Fermi surface. RPA results have $\Gamma = 1, G(q) = 0$. Vashishta–Singwi results listed as VS have $\Gamma = 1/(1+\chi G)$, with their choice for $G(q)$. From Mahan and Sernelius.

integrand is badly behaved at the origin. The dielectric function $\varepsilon(q, i\omega)$ is singular in the region where the wave vector and frequency are both very small. This region of integration must be treated with great care. Table 4 gives some results from Mahan and Sernelius. The top two rows list the RPA results, while the lower two rows list results found using the Vashishta-Singwi form for $G(q)$. The RPA results for the self-energy are larger in magnitude. Band narrowing is defined here as the difference between the values for $\Re\Sigma(p)$ at $p = 0$ and at $p = k_F$. There is not much difference in band narrowing as predicted by RPA and by using VS form for the Hubbard correction.

Electron density is given in terms of the dimensionless quantity $r_s a_B = (3/4\pi n)^{1/3}$, where the Bohr radius is a_B. Aluminum has $r_s = 2.07$. The table shows little band narrowing at $r_s = 2$, which is close to aluminum. Sodium has $r_s = 3.96$. The table shows the band narrowing at $r_s = 4$ is -0.026 Ry which is the value of -0.35 eV listed in Table 2. Similar results are also found using other choices for $G(q)$. These numerical results are shown graphically in figure 8. The solid line includes vertex corrections and local field factors. That curve is very similar to RPA. The curve labeled "$\Gamma = 1$" is the GW approximation, which gives a poor result.

The 'GW' approximation is defined as setting the factor $\Gamma = 1$ in the numerator of (20). If one sets $\Gamma = 1$ everywhere, one has RPA. The GW approximation uses the vertex correction in the dielectric function, but not in the numerator of the self-energy. It is widely used in calculations of band gaps in semiconductors[43–44]. Our own calculations show that the GW ap-

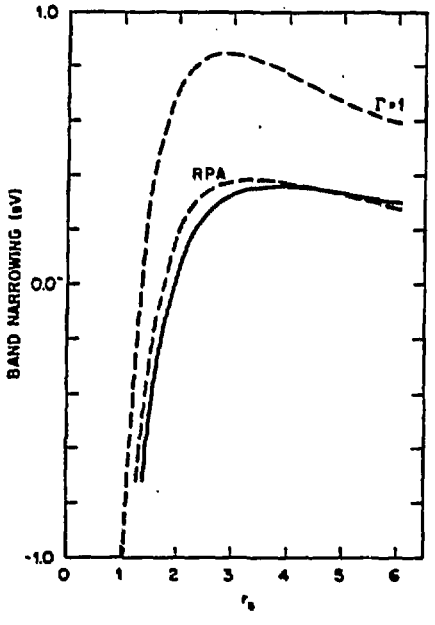

Figure 8: Band narrowing as a function of density. The solid line includes vertex corrections. The RPA curve is similar. The dashed curve labeled "$\Gamma = 1$" is the GW approximation, which is poor. From Mahan and Sernelius.

proximation is very poor in free electron metals. The top curve shows the predicted band narrowing obtained using the GW approximation. This result is very different than found when vertex corrections are included correctly. Lyo and Plummer published a calculation of band narrowing. They made two approximations: the GW approximation, and the plasmon-pole approximation. The latter approximates the spectral function $B(q,\omega)$ by a delta function. We have tested both of these approximations, and found they are poor. The Lyo-Plummer results are very different than what we found by making neither approximation.

The kind of band narrowing we are calculating are corrections to the effective mass. If there is a self-energy difference between $p = 0$ and $p = k_F$, and if it is a smooth function of wave vector p, then it can be represented as a contribution to the effective mass $m*$. Our conclusion is that electron-electron interactions do not contribute anything to the effective mass beyond what was found twenty-five years ago by Hedin and Rice.

Are there other ways in which electron-electron calculations can contribute to the effective mass? At the present time, we are doing calculations to test another possibility. It is the change in the chemical potential due to band tails. The band tails are also calculated using electron-electron interactions. Figure 9 shows the density of states calculated by Hedin and Lundqvist[45] for the homogeneous electron gas with $r_s = 2$. There are two notable features to this result: (i) the long tail at low energy. Our calculations show that 20-25% of the band electrons are in this band tail; (ii) In the band itself, the density of states is very much reduced. This curve is calculated from the derinition of the density of states

$$\rho(\omega) = 2\int \frac{d^3p}{(2\pi)^3} A(p,\omega), \tag{30}$$

$$= -\frac{1}{\pi^2}\int p^2 dp \frac{\Im\Sigma(p,\omega)}{(\epsilon_p - \omega - \Re\Sigma(p,\omega))^2 + (\Im\Sigma(p,\omega))^2}. \tag{31}$$

After the calculation of the self-energy, one has to integrate over all wave vector in order to find the density of states. The chemical potential μ is determined by the requirement that the number of particles is fixed

$$n = \int d\omega \rho(\omega)\Theta(\mu - \omega). \tag{32}$$

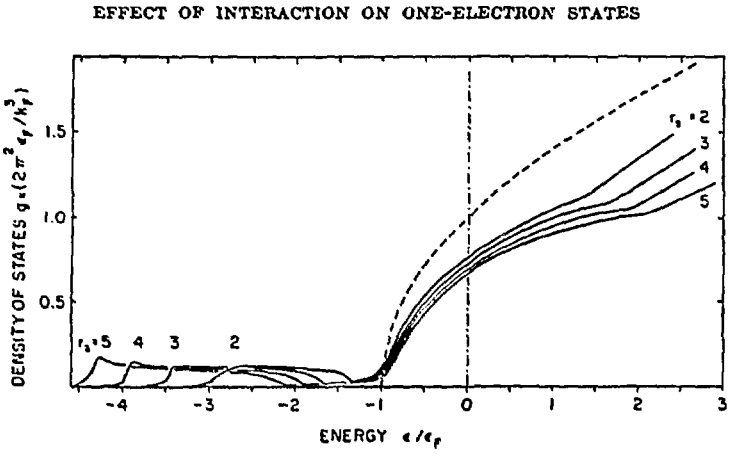

Figure 9: Density of states from electron-electron interactions as calculated by Hedin and Lundqvist.

The long bandtail in the density of states is caused by $\Im\Sigma(p,\omega)$, which is nonzero for very large values of negative energy ω.

The position of the chemical potential determines the band width. A smaller band width occurs if μ is lowered. Does the long band tail cause the chemical potential to be lowered? It has a tendency to lower it, and the depleted density of states in the main band causes the chemical potential to be raised. Which of these two opposing tendencies is larger? We have undertaken a detailed calculation of the density of states, in order to answer this question. Calculations are underway. No results on band narrowing are present at this time. However, we have determined one result. The inclusion of the local field factor $G(q)$ has no effect on $\rho(\omega)$. We get the same curve in RPA that we do with GV choice of the local field correction.

Research support acknowledged by the National Science Foundation through grant number DMR-9015771, from the University of Tennessee, and by the U.S. Department of Energy through contract DE-AC05-84OR21400 administered by Martin Marietta Energy Systems Inc.

6 References

1. A. Sommerfeld, Z. Physik **47**, 1 (1928)
2. H.W.B. Skinner, Phil. Trans. Roy. Soc. **A239**, 95 (1940)
   H.M. O'Brian and H.W.B. Skinner, Phys. Rev. **45**, 370 (1934)
3. F. Seitz, *Modern Theory of Solids* (McGraw Hill, 1940) pg 436
4. J. Hubbard, Proc. Roy. Soc. (London) **A 243**, 336 (1957)
5. K.S. Singwi, M.P. Tosi, R. H. Land, and A. Sölander, Phys. Rev. **176**, 589 (1968)
6. K.S. Singwi, A. Sjölander, M.P. Tosi, and R.H. Land, Phys. Rev. B **1**, 1044 (1970)
7. P. Vashishta and K.S. Singwi, Phys. Rev. B **6**, 875 (1972)
8. K.S. Singwi and M.P. Tosi, *Solid State Phys.* **36**, 177 (1981)
9. D. Ceperley, Phys. Rev. B **13**, 3126 (1978)
10. J.D. Talman , Phys. Rev. A **13**, 1200 (1974)
11. L. Hedin, Phys. Rev. **139**, A796 (1965)
12. T.M. Rice, Ann. Phys. **31**, 100 (1965)
13. E. Jensen and E.W. Plummer, Phys. Rev. Lett. **55**, 1912 (1985)
14. H.J. Levinson et al., Phys. Rev. B **27**, 727 (1983)
15. E. Jensen et al., Phys. Rev. B **30**, 5500 (1984)
16. I.W. Lyo and E.W. Plummer, Phys. Rev. Lett. **60**, 1558 (1988)
17. B.S. Itchkawitz, I.W. Lyo, E.W. Plummer, Phys. Rev. B **41**, 8075 (1990)
18. U.S. Whang, E.T. Arakawa, and T.A. Callcott, Phys. Rev. B **6**, 2109 (1972)

39. S.S. Schweber, *Introduction to Relativistic Quantum Field Theory* (Harper Row, 1961) pg 613

40. C.A. Kukkonnen and A.W. Overhauser, Phys. Rev. B **20**, 550 (1979)

41. C.S. Ting, T.K. Lee, and J.J. Quinn, Phys. Rev. Lett. **34**, 870 (1975)

42. C. Petrillo and F. Sacchetti, Phys. Rev. B **38**, 3834 (1988)

43. J.E. Northrup, M.S. Hybertsen and S.G. Louie, Phys. Rev. Lett. **59**, 819 (1987)

44. M.P. Surh, J.E. Northrup, and S.G. Louie, Phys. Rev. B. **38**, 5976 (1988)

45. L. Hedin and S. Lundqvist, *Solid State Phys.* **23**, 1 (1969)

19. K.W.K. Shung and G.D. Mahan, Phys. Rev. Lett. **57**, 1076 (1986)

20. K.W.K. Shung and G.D. Mahan, Phys. Rev. B **38**, 3856 (1988)

21. K.W.K. Shung, B.E. Sernelius, and G.D. Mahan, Phys. Rev. B **36**, 4499 (1987)

22. G.D. Mahan and B.E. Sernelius, Phys. Rev. Lett. **62**, 2718 (1989)

23. G.D. Mahan, Phys. Rev. B **2**, 4334 (1970)

24. H. Neddermeyer, Z. Physik **271**, 329 (1974)

25. P. Livins and S.E. Schnatterly, Phys. Rev. B **37**, 6731 (1988)

26. G.D. Mahan, *Many-Particle Physics*, Second Ed. (Plenum, 1990) Ch.8

27. D.R. Penn, S.M. Girvin, and G.D. Mahan, Phys. Rev. B **24**, 6971 (1981)

28. G.D. Mahan, Phys. Rev. B **25**, 5021 (1982)

29. K. Ohtaka and Y. Tanabe, Revs. Mod. Phys. **62**, 929 (1991)

30. P. Minnhagen, Phys. Lett. **56A**, 327 (1976)

31. P.H. Citrin et al., Phys. Rev. Lett. **61**, 1021 (1988)

32. T.A. Callcott, D.L. Ederer, Phys. Rev. Lett. **62**, 1322 (1989)
Phys. Rev. B **18**, 6622 (1978)

33. P.B. Allen, M.L. Cohen, D.R. Penn, Phys. Rev. B **38**, 2513 (1988)

34. D. Pines, *Many-Body Problem* (Benjamin, 1962)

35. G. Vignale and K.S. Singwi, Phys. Rev. B **32**, 2156 (1985)

36. N. Iwamoto, E. Krotscheck, and D. Pines, Phys. Rev. B **29**, 3936 (1984)

37. G.D. Santoro and G. Giuliani, Phys. Rev. B**37**, 4813 (1988)

38. D.J. W. Geldart and S.H. Vosko, Can. J. Phys. **44**, 2137 (1966)