# ELECTRON–ELECTRON INTERACTIONS IN DISORDERED SYSTEMS

*Volume editors*

## A.L. EFROS
*Leningrad, USSR*

## M. POLLAK
*Riverside, California, USA*

MODERN PROBLEMS IN CONDENSED MATTER SCIENCES

*Oh, how many of them there*
*are in the fields!*
*But each flowers in its*
*own way—*
*In this is the highest achievement*
*of a flower!*

*Matsuo Bashō*
*1644–1694*

# PREFACE TO THE SERIES

"Modern Problems in Condensed Matter Sciences" is a series of contributed volumes and monographs on condensed matter science that is published by North-Holland Publishing Company. This vast area of physics is developing rapidly at the present time, and the numerous fundamental results in it define to a significant degree the face of contemporary science. This being so, it is clear that the most important results and directions for future developments can only be covered by an international group of authors working in cooperation.

Both Soviet and Western scholars are taking part in the series, and each contributed volume has, correspondingly, two editors. Furthermore, it is intended that the volumes in the series will be published subsequently in Russian by the publishing house "Nauka".

The idea for the series and for its present structure was born during discussions that took place in the USSR and the USA between the former President of North-Holland Publishing Company, Drs. W.H. Wimmers, and the General Editors.

The establishment of this series of books, which should become a distinguished encyclopedia of condensed matter science, is not the only important outcome of these discussions. A significant development is also the emergence of a rather interesting and fruitful form of collaboration among scholars from different countries. We are deeply convinced that such international collaboration in the spheres of science and art, as well as other socially useful spheres of human activity, will assist in the establishment of a climate of confidence and peace.

The General Editors of the Series,

V.M. Agranovich

A.A. Maradudin

# INTRODUCTION

I am delighted to write a short introduction to this volume. I congratulate the editors on having brought together contributions from so many of the countries in which scientists have been at the forefront of this field of research. I have myself been involved in it since 1949, and shall study this volume with the greatest interest.

I suppose the most important single step in the study of Coulomb interactions in solids was the demonstration by Landau that, in spite of these interactions, the Fermi surface in $k$-space is a real physical quantity and at zero temperature is sharp. So of course is the sharp cut-off at the Fermi energy, which was shown even earlier in 1934 by Jones, Mott and Skinner. This is true also for a degenerate electron gas in a non-crystalline field. As regards the problem of electrons in these fields, even without interactions, I doubt if it is fully understood even yet. Such understanding as we have starts with P.W. Anderson's famous paper of 1958 on the "Absence of Diffusion in Certain Random Lattices", which shows not unexpectedly that disorder must lead to traps in a conduction band but, more surprisingly that traps with a continuous range of energies can prevent diffusion. In the light of Anderson's concept of localization, one can divide problems into those of the conduction band of an amorphous semiconductor, such as hydrogenated amorphous silicon, in which Coulomb interactions are not important, and those of a degenerate gas, for instance in an impurity band, where they are. Even the former are by no means understood; in real semiconductors there is no agreed calculation of the position of the mobility edge or of the shape of the tail of localized states; the sign of the Hall effect remains to be explained and there is lack of agreement on whether polarons are formed.

For the degenerate electron gas, a puzzle until recently has been the status – if any – of the author's "minimum metallic conductivity", given by

$$\sigma_{\min} \simeq 0.025\, e^2/\hbar a,$$

where in a compensated impurity band $a$ is the distance between centres. There was much experimental evidence for its existence, but in 1979 arguments based on scaling theory claimed that $\sigma(T = 0)$ would always go continuously to zero as the Fermi energy $E_F$ tends to the mobility edge $E_c$, and there is now

much experimental evidence that this is so, if experiments are carried out at a low enough temperature. Recently Mott (1984) has shown that this can be deduced from the Kubo–Greenwood formula. A minimum metallic conductivity will only exist if a transition is induced by a magnetic field, and is then

$$\sim 0.025\, e^2/\hbar L,$$

where $L = L_H = (c\hbar/He)^{1/2}$ if $L_H > a$, and $L = a$ otherwise. The author's "minimum metallic conductivity" should be observed, then, for magnetic fields such that $L_H < a$ at the transition, as is the case in the work of Long and Pepper (1984) and Dubois et al. (1984) on doped InP, using temperatures down to 40 mK.

Turning now to interactions, the intra-atomic interaction $\langle e^2/r_{12} \rangle$, called the Hubbard $U$, is responsible for a metal–insulator transition in a crystalline array of one-electron centres – or more generally centres with a magnetic moment. This is what is often called a "Mott transition". The most interesting point about it is that, in crystalline materials, it is predicted to be a first-order transition; in a rigid lattice the number of free electrons should jump discontinuously from zero to a finite value. This discontinuity may not exist in strongly disordered systems (Mott 1978) – but probably does in, for instance, metal–ammonia solutions, where a solubility gap appears to be a consequence of the transitions. But whether this is so in uncompensated doped semiconductors is very doubtful; there is no evidence for it at all.

Doped semiconductors do of course provide the ideal field for the exploration of interactions in degenerate gases, and in this the work of Gordon Thomas and co-workers reported in this book is outstanding. They were the first to show that, at low enough temperatures, the conductivity would go continuously to zero if the transition occurred through change of concentration or stress (in contrast, in our view, to magnetic fields). Now they find that $\sigma$ is proportional to $(n_c - n)^s$ with $s = \frac{1}{2}$, instead of $s = 1$ as predicted by a theory without interactions and observed in other systems such as a Si–Nb. At the time of writing, several theorists are attempting to explain this in terms of interactions; can it be related to the change in the density of states near the Fermi level, as $|E - E_F|^{1/2}$, first predicted by Altshuler and Aronov (1979)?

The transition in doped semiconductors is usually supposed to take place in an impurity band – but for Si : P this must be doubted, because the electronic specific heat shows little deviation from the value calculated for free electrons in the silicon conduction band. We have to ask, then, can an Anderson transition take place without a drop in the density of states resulting from disorder? Perhaps this is only possible in a many-valley band.

In the limit of low concentrations of dopant, there is of course no doubt that conduction is in an impurity band and it is here, for hopping conduction, that interactions through the "Coulomb gap" have their most striking effect. It is

good to see contributions from the two editors, pioneer workers on this subject, setting out so fully the present position.

Sir Nevill Mott
Cavendish Laboratory
Cambridge, UK

## *References*

Dubois, H., G. Biskupski, I.L. Wojkiewicz, H. Briggs and G. Remenÿl, 1984, J. Phys. **C17**, L411.
Jones, H., N.F. Mott and H.W.B. Skinner, 1934, Phys. Rev. **45**, 378.
Long, A.F., and M. Pepper, 1984, J. Phys. **C17**, 3391.
Mott, N.F., 1978, Phil. Mag. **B37**, 377.
Mott, N.F., 1984, Phil. Mag. **B49**, L75.

# CONTENTS

# Electron–Electron Interaction In Disordered Conductors

B.L. ALTSHULER and A.G. ARONOV

*Leningrad Nuclear Physics Institute*
*Gatchina, Leningrad, 188350*
*USSR*

# Contents

3

## 1. Introduction

The theory of transport phenomena based on the classical equation assumes that between scattering events the electron moves along a classical trajectory. This is valid, however, only provided interference of scattering from different centers can be neglected. The criterion for the validity of such an approach is the smallness of the electron wavelength $\lambda$ compared to the mean free path $l$, as pointed out by Ioffe and Regel (1960), Mott (1961), and Mott and Davis (1971). If this condition is not satisfied, then, as shown by Anderson (1958), the electron states are localized, the zero temperature conductivity being zero.

The momentum relaxation of electrons in disordered conductors at sufficiently low temperatures is governed by their elastic scattering from impurities and structural defects. Even if $\lambda \ll l$, quantum corrections to the transport effects become essential, since it is these corrections that are responsible for the nontrivial dependences on temperature, external fields and sample size (Abrahams et al. 1979, Anderson et al. 1979, Gor'kov et al. 1979). These corrections are applicable even in the case of an ideal gas of noninteracting electrons.

Electron–electron interaction in disordered metal systems also results in nontrivial corrections which are small in the $\lambda/l$ parameter and depend on temperature, external fields and sample size. However, in contrast to the case of noninteracting electrons, such corrections have to be applied not only to the transport, but also to thermodynamic quantities and to the electron density of states at the Fermi level as well (Altshuler and Aronov 1979a, b, c, Altshuler et al. 1980a, b, Altshuler et al. 1982c, d, Fukuyama 1980a).

This effect of the electron–electron interaction in disordered systems makes it drastically different from that for pure metals, where the interaction at low temperatures manifests itself only in renormalization of the electron spectral parameters.

Progress in the theory of disordered metals has provided clues to a number of phenomena which have been defying explanation for a long time.

Among them is the negative magnetoresistance in semiconductors discovered more than 20 years ago, zero-bias anomalies in the tunneling conductivity, the minimum in the temperature dependence of the resistivity of disordered metals, which are not associated with the Kondo effect, and some other phenomena. At the same time, investigation of the quantum corrections to the transport phenomena has offered a unique possibility to shed light on the microscopic relaxation times in metals and their temperature dependence,

among them the inelastic collision time, the phase relaxation time, the spin–orbit and spin–spin scattering times (Bergmann 1982a, b, c, Gershenzon and Gubankov 1981, Gershenzon et al. 1982a, b).

The present review deals with the investigation of the effect of electron–electron interaction on the properties of disordered conductors in the metallic conduction domain. The theory developed in the recent years and outlined in this review is applicable to dirty and amorphous metals, metal glasses and doped semiconductors. Similar to the case of quantum corrections to the kinetic effects in a noninteracting electron gas, the interaction effects are most clearly pronounced in low-dimensionality systems, i.e., in the MOS structures in semiconductors, in thin films and wires.

The purpose of the present survey is to give a consistent overview of the effect of electron–electron interaction on the properties of disordered conductors. Therefore, despite the wealth of available experimental data on the quantum corrections in disordered conductors, we will present here only a few of them just to illustrate the major achievements of the theory, rather than attempt to give a comprehensive critical analysis of the experimental situation in this area.

The review is constructed in the following way.

In section 2, the major results of the application of quantum corrections to the theory of a noninteracting electron gas are studied qualitatively, the principal phenomena are considered, and the effect of external fields is discussed. In the same section (subsection 2.2) the conventional diagrammatic technique is outlined, and subsequently the principal quantities used are introduced.

Section 3 reports on a study of the effect of electron–electron interaction on the one-particle density of states and the thermodynamics of disordered conductors. It is shown, in particular, that particle repulsion produces a minimum in the density of states at the Fermi energy. It is this minimum that is responsible for the anomalous maximum of tunneling resistivity at zero bias. The effective electron interaction for the case where the gas approximation fails is also discussed. Quantum corrections to such thermodynamic quantities as the heat capacity and compressibility of the electron gas are considered in the concluding part of this section.

Section 4 deals with an investigation of electron–electron collisions in disordered metals. The presence of a large number of impurities or defects changes the classical result of Landau concerning the temperature dependence of the inelastic electron–electron collision time. It is found to be different for bulk samples, films and thin wires.

The electron–electron collisions result in relaxation not only of the occupancy, but of the quasiparticle wave function phase relaxation as well.

When studying the phase relaxation time $\tau_\varphi$ of the wave function, we

encounter an unusual situation in the kinetics of normal metals, when the quasiclassical description of this relaxation process is not applicable, since the characteristic energy transfer turns out to be of the same order as the inverse relaxation time multiplied by Planck's constant. Because of the unusual nature of this situation, we have given here a detailed mathematical derivation of the basic results since this situation, in our opinion, is not so unique as not to be met more than once in the kinetics of normal metals.

Section 5 is devoted to the effect of electron–electron interaction on transport phenomena in disordered conductors, as well as on the temperature and frequency dependence of the conductivity and the Hall effect. Just as in the preceding sections, we have tried here to sum up the various contributions to these effects providing the same functional dependence on the main quantities, such as temperature, frequency, degree of disorder, while originating from different interaction effects.

In section 6 we treat the effect of magnetic field and spin scattering on the phenomena in question. The magnetic field, just as the spin-flip electron scattering (both of the spin–orbit and spin–spin type) provide new degrees of freedom and thus permit one to separate the various contributions to the quantum effects. Indeed, the interaction effects begin, for instance, to depend on magnetic field at higher fields than is the case with the quantum corrections to the conductivity of noninteracting electrons. Spin scattering also acts in a different way, its effect depending on whether it is of the spin–spin or spin–orbit type. Also, the various contributions are differently affected by the rate of this scattering.

In the same section, the behavior of the magnetic susceptibility of disordered metals is analyzed. It is found that the quantum correction to the orbital magnetic susceptibility originating from electron interaction cannot be small (Aslamasov and Larkin 1974), and can govern not only the temperature and field dependence of the magnetic susceptibility, thus imitating the localized spin susceptibility, but also its magnitude in low-dimensionality systems.

Section 7 deals with an investigation of the effect of anisotropy and intervalley transitions on the interaction effects in disordered semiconductors.

Despite the obvious achievements in the theory of disordered conductors, the very important question of how the metal–insulator transition occurs with a change of disorder, or of the Fermi energy, remains unanswered. Considerable progress in this domain has been recently achieved without including electron–electron interaction following the work of Abrahams et al. (1979) who showed that in three dimensions the transition is essentially of the second kind, i.e., without a jump in conductivity, whereas in two- and one-dimensional systems the electrons are always localized. Wegner (1979) and Efetov et al. (1980) showed that the field-theoretical model describing this transition is renormalizable, and thus supported the main results obtained by Abrahams et al. (1979).

The effect of electron–electron interaction on the transition proper has not, however, been clarified. McMillan (1981) developed a phenomenological scaling theory for the metal–insulator transition; however, some of its major conclusions have been criticized by Castellani et al. (1982) and Finkelshtein (1983). The latter demonstrated also renormalizability of the theory of the metal–insulator transition taking into account the interactions.

In section 9 the metal–insulator phase transition is considered including interactions in the presence of paramagnetic impurities. It is shown that this phase transition does not differ qualitatively from Anderson's transition for the case of a noninteracting electron gas without spin scattering.

Those parts of the mathematics involved which are not necessary for the understanding of the subject are transferred to the Appendices.

## 2. Noninteracting electrons in disordered metal systems

### 2.1. Quantum corrections to the conductivity of noninteracting electrons

The residual resistance in conductors is determined by the elastic scattering of electrons from impurities and static lattice defects. In "good" conductors, the wavelength of the electron $\lambda = 2\pi\hbar/p$ ($p$ is the electron momentum) is much less than its mean free path $l$. Between collisions the electron moves as a free particle and can be described in a quasiclassical way. This description of electronic behavior yields the well-known Drude expression for the conductivity

$$\sigma = \frac{e^2\tau}{m}N,\tag{2.1}$$

where $N$ is the electron concentration, $m$ the effective mass, $\tau$ the time between collisions, $e$ the electronic charge. As shown by Anderson (1958), at a sufficiently high impurity concentration when $\lambda > l$, the electron states become localized and no longer contribute to conduction. The transition from delocalized to localized states observed to occur with increasing degree of disorder is called Anderson's transition.

However, even in the domain of high metallic conductivity when $\lambda \ll l$, quantum corrections can be added to the classical expression of Drude. In this section we will discuss the physical meaning of these corrections, their magnitude and some phenomena associated with them.

Consider a good conductor ($\lambda \ll l$). To get from point A to point B (fig. 1), a particle can move along different paths. The total probability $w$ for a transfer from point A to point B can be obtained by squaring the modulus of the sum of all amplitudes of the probability for a particle to pass along all possible paths, i.e.

$$w = \left|\sum_i A_i\right|^2 = \sum_i |A_i|^2 + \sum_{i\neq j} A_i A_j^*.\tag{2.2}$$

The first term in eq. (2.2) represents the sum of probabilities for the particle to pass by any way, and the second, the interference of various amplitudes. For most paths the interference is not essential since their lengths differ strongly, and hence also the phases of the wave functions should differ substantially on these paths. Therefore, when summing over all paths, the mean value of the interference term will vanish because of its oscillating nature. There are, however, paths of a specific kind, namely, self-intersecting paths (fig. 1). Each path of this kind can be assigned a pair of amplitudes, $A_1$ and $A_2$, corresponding to the passage of the loop clock- and counterclockwise. These two amplitudes are coherent and thus interference of these waves may not be neglected. As a result, the probability to find a particle at point O becomes

$$|A_1|^2 + |A_2|^2 + 2\operatorname{Re} A_1 A_2^* = 4|A_1|^2,$$

i.e., twice the figure we would obtain if we summed up the probabilities neglecting interference. On the other hand, an enhanced probability to find a particle at the point it has left means a reduced probability to find it at point B (the point of observation), i.e., a decrease of conductivity or increase of resistivity, induced by interference. Neglect of interference corresponds to a classical description of the electrons (Boltzmann equation), and its inclusion corresponds to quantum corrections to the conductivity.

Estimate now the magnitude of the corrections to the conductivity. Because of the collisions with impurities, the electron paths follow a random walk pattern. The interfering amplitudes correspond to the paths contained within a ray tube of cross section $\lambda^2$. The relative magnitude of the quantum correction to conductivity, $\delta\sigma/\sigma$, which, as we have just seen, is negative, is proportional to the probability for such a ray tube to intersect itself in a classical diffusion motion

$$\frac{\delta\sigma}{\sigma} \sim -\int_\tau^{\tau_\varphi} \frac{v\lambda^2 \, \mathrm{d}t}{(Dt)^{3/2}}, \qquad (2.3)$$

where $v$ is the particle velocity and $D$ the diffusion coefficient. Integration in eq. (2.3) is performed within the limits $\tau < t < \tau_\varphi$, where $\tau_\varphi$ is the phase
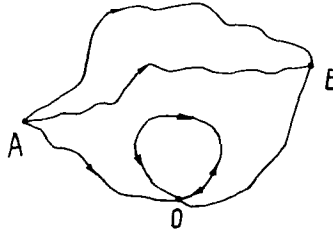


Fig. 1. Various paths for a particle to move from point A to point B. O = point of path self-crossing.

relaxation time associated with inelastic or spin-flip scattering. (Later we shall discuss the physical meaning of $\tau_\varphi$.) Thus the quantum correction to conductivity acquires the form (Anderson et al. 1979, Gor'kov et al. 1979)

$$\delta\sigma \sim \frac{e^2}{\hbar L_\varphi} + \text{const.}, \tag{2.4}$$

where $L_\varphi = \sqrt{D\tau_\varphi}$. As seen from eq. (2.4), although the quantum correction to conductivity is small in the parameter $\hbar/pl \ll 1$, nevertheless it brings about a nontrivial temperature (since $\tau_\varphi \sim T^{-p}$) or frequency relationship (at $\Omega\tau_\varphi \gg 1$ the upper limit of integration in eq. (2.3) has to be replaced by $1/\Omega$ and, hence, $L_\varphi$ in eq. (2.4) is replaced by $L_\Omega = \sqrt{D/\Omega}$). It should be emphasized that the interference effects produce a growth of resistivity as the temperature decreases.

If the transverse size of a film or wire $a \ll L_\varphi$, the particle will be able to diffuse many times the distance from one wall to another in time $\tau_\varphi$, so that the probability to find it at any point along the transverse dimension will be the same. Hence for the quantum correction to conductivity we obtain

$$\frac{\delta\sigma}{\sigma} \sim -\int_\tau^{\tau_\varphi} \frac{v\lambda^2 \, \mathrm{d}t}{(Dt)^{d/2}} \frac{1}{a^{3-d}}, \tag{2.3a}$$

where $d$ is the effective dimensionality of the sample ($d = 2$ for a film and $d = 1$ for a wire). On integration we come to

$$\begin{aligned} \delta\sigma_d &\sim -\frac{e^2}{\hbar} \ln L_\varphi/l, \quad d = 2, \\ &\quad -\frac{e^2}{\hbar} L_\varphi, \qquad d = 1. \end{aligned} \tag{2.4a}$$

Here $\sigma_d = \sigma a^{3-d}$. (For a film $\sigma_2$ is the conductivity of a square sample, for a wire $\sigma_1$ refers to a sample of unit length.) Expression (2.4a) is valid at $d = 2$ also for a purely two-dimensional case, e.g., MOSFET devices, where the motion of particles in the transverse direction is quantized, making only planar diffusion possible.

One may thus conclude that the cross-over from one dimensionality to another occurs when the transverse size of a sample becomes equal to $L_\varphi$.

Consider now the physical meaning of $\tau_\varphi$ following Altshuler et al. (1981a, b). $\tau_\varphi$ is the time in which the wave function retains its coherence. It should be stressed that this time is neither the out-relaxation time (by which we understand here the relaxation time associated with the scattering-induced transfer of a particle to any quantum state) nor the energy relaxation time of the distribution function. To illustrate this point, consider the case of quasielastic electron scattering by acoustic phonons. Such a relaxation mechanism can operate in degenerate semiconductors at not too high electron concentrations or in metals, when the temperature is about the Debye temperature.