LOCALIZATION AND LOW TEMPERATURE
TRANSPORT IN DISORDERED ONE-DIMENSIONAL SYSTEMS

by

ALFRED DOUGLAS STONE
B.A., Harvard College (1976)
M.A., Balliol College, Oxford (1978)

SUBMITTED TO THE DEPARTMENT OF
PHYSICS IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS OF THE
DEGREE OF
DOCTORATE IN PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1983

© Massachusetts Institute of Technology 1982

Signature of Author

Department of Physics

Certified by

Signature redacted

John D. Joannopoulos
Thesis Supervisor

Accepted by

Signature redacted

George F. Koster
Head, Ph.D. Committee
Acknowledgements

One of the nice things about physics as a field is that it is very conducive to collaboration. I would like to thank the many people who helped me or collaborated with me on this work. I thank D. Vanderbilt, B. Kleiman, D. H. Lee, M. Kastner, D. Monroe, E. Abrahams, and P. W. Anderson for helpful discussions on various aspects of this work. I am very grateful to T. Goodwillie for his mathematical assistance, and to J. Bather, H. Chernoff and W. DuMouchel of the MIT statistics center, for the donation of their statistical wisdom. I am also very grateful to Patrick A. Lee for his guidance in the derivation of the Landauer formula, for his helpful criticism of my original approach for calculating inelastic lifetimes, and for my catchy opening paragraphs, which I shamelessly "borrowed" from him.

I am also, of course, greatly indebted to my collaborators, D. Allan, D. J. Chadi, and D. J. Thouless. Mere words cannot express my gratitude to Denise Borsuk for her excellent typing of this manuscript. And finally, I would particularly like to thank my supervisor, John D. Joannopoulos. Not only has John been a source of physical insight into almost all aspects of this work, he has also been extremely generous with his time and personally supportive while I was making the difficult transition from student to researcher. I greatly appreciate all his efforts on my behalf.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>1</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>2</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>3</td>
</tr>
<tr>
<td>I Abstract</td>
<td>4</td>
</tr>
<tr>
<td>I Introduction</td>
<td>5</td>
</tr>
<tr>
<td>II The Landauer Conductance Formula</td>
<td>15</td>
</tr>
<tr>
<td>III Exact Scaling Law for Average Resistance of the 1d Anderson Model</td>
<td>28</td>
</tr>
<tr>
<td>IV Exact Solution for Off-Diagonal Disorder at Bandcenter</td>
<td>42</td>
</tr>
<tr>
<td>V Validity of the Random Phase Hypothesis</td>
<td>53</td>
</tr>
<tr>
<td>VI Strange Statistics</td>
<td>69</td>
</tr>
<tr>
<td>VII Conduction at Finite Temperature: The Role of Statistical Fluctuations</td>
<td>84</td>
</tr>
<tr>
<td>VIII Inelastic Scattering Models for Weak Localization Effects in Thin Wires</td>
<td>97</td>
</tr>
<tr>
<td>IX Conclusions</td>
<td>132</td>
</tr>
<tr>
<td>References</td>
<td>135</td>
</tr>
<tr>
<td>Figure Captions</td>
<td>144</td>
</tr>
</tbody>
</table>
LOCALIZATION AND LOW TEMPERATURE TRANSPORT IN DISORDERED ONE-DIMENSIONAL SYSTEMS

by

ALFRED DOUGLAS STONE

Submitted to the Department of Physics on January 7, 1983 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Physics

ABSTRACT

We consider the problem of DC electrical conduction in disordered one-dimensional solids. At zero temperature, the system's eigenstates are expected to be localized in space, causing the resistance to increase exponentially with length. We derive rigorously a general expression for the DC conductance of such a system, the Landauer formula. Using this formula, we calculate exact expressions for moments of the resistance in a generalized one-dimensional Anderson model, and solve for the probability distribution of the resistance in a special case. The unusual statistical behavior of the average resistance is explained in terms of the properties of lognormal distributions. A new technique for calculating meaningful averages is developed and shown to agree with numerical results. The role of statistical fluctuations in experiments at finite temperatures is examined and shown to be negligible for the range of system parameters presently accessible. A calculation of the inelastic scattering length, which determines the size of the localization effect at finite temperature in a thin wire, is given. The remaining discrepancies between theory and experiment are discussed.

Thesis Supervisor: Professor John D. Joannopoulos
Title: Associate Professor of Physics
Chapter I. Introduction

There is a standard description of electrical conduction in metals which every physics student learns in his first course on solid state physics. Metals are solids with a partially-filled energy band. The electrons in this band can be excited by arbitrarily small amounts of energy (unlike those in semi-conductors or insulators). Thus, they effectively form an electron gas, which may be treated as non-interacting by using a fermi liquid type of approximation. The conductivity of this gas can be obtained simply by assuming that the electron scatters in a time $\tau$ due to both thermal and intrinsic defects in the solid, and then applying semi-classical arguments to obtain the Ohmic conductivity $\sigma = ne^2\tau/m$. As $T \to 0$, $\sigma$ increases monotonically to a finite, non-zero value $\sigma_0$ at $T = 0$, the signature of metallic conduction.

There is almost nothing right about this description for a wide class of "metallic" solids. That is the exciting result of modern investigations of low temperature conduction in disordered solids. It is now known that many disordered "metals" are in fact insulators as $T \to 0$, with zero DC conductance. There is also an emerging consensus that many-body interactions play a crucial role in determining the nature of the metal-insulator transition. This thesis will deal primarily with only one of the mistaken aspects of the conventional description, the effect of a disordered static potential on non-interacting electrons, which does not simply introduce a finite elastic scattering rate into the electron gas. However, we will also discuss the role of interactions, particularly with regard to their effects at finite temperature.
In fact, meaningful finite temperature results can only be obtained by including interactions; a thermal average over non-interacting electrons misses a characteristic feature of transport in disordered metals: at low temperatures thermal scattering assists transport instead of impeding it.

The study of a system of non-interacting electrons in a random, static potential at zero temperature has become known as the localization problem. The name arising from the fact that, under very general circumstances which will be discussed in detail below, the eigenstates of such a random potential are peaked at a particular point in space and have an envelope which decays exponentially with the distance away from the maximum. This is of course to be contrasted with the Bloch functions of a perfect crystal whose moduli are periodic functions which extend throughout the crystal. The first argument that randomness should localize the electron is due to Anderson. He considered a system of isolated atoms each with a single bound state of random binding energy \( \epsilon \) distributed uniformly over a width \( W \). He then introduced a constant hopping interaction \( V \) between nearest neighbor atoms and asked for the probability that an electron initially on a site \( i \) at \( t = 0 \), will still be there as \( t \to \infty \). If this probability goes to zero, the eigenstates at the energy \( \epsilon_i \) are extended; if not they are localized. He first showed that, as you would expect, this probability can only be zero if the perturbation expansion giving corrections to the localized state \( |\epsilon_i\rangle \) diverges. He then showed that this expansion did diverge, with probability unity (i.e. for almost all choice of the \( \{\epsilon_i\} \)), when the ratio \( V/W \) became sufficiently large. This established the possibility
of a quantum-mechanical phase transition at zero temperature, in which for a given energy, the qualitative character of the eigenstates at that energy would be changed as a function of the degree of disorder in the system. This transition from extended to localized states is now known as an Anderson transition.

Mott\(^2\) built on these ideas by noting that extended and localized states cannot exist at the same energy, since any perturbation would then mix them, creating two extended states. Therefore, the electronic density of states in such systems should be divided into a region of extended states surrounded by regions of localized states, which meet at a particular energy called the mobility edge. If the system's fermi energy lies in the extended region it should behave like a conventional metal as \(T \to 0\), whereas if it lies in the localized region conduction could only occur by thermal excitation to the mobility edge, and at \(T = 0\) it would be an insulator. These concepts shifted the focus of research away from the nature of the eigenstates in disordered systems and towards the study of discontinuities in the behavior of the zero temperature transport properties as a function of the degree of disorder and the fermi energy.

The next significant advance in localization theory came by considering the behavior of the zero temperature conductance \(G\), as a function of the system size. For a conventional metal at \(T = 0\) this dependence was well-known from Ohm's law, which says that for a system of volume \(L^d\)

\[
G = \sigma_0 L^{d-2}
\]  

(1-1)
where $\sigma_0$ is the zero temperature Ohmic conductivity. In the disordered metal Eq. (1) is not expected to be valid for all degrees of disorders and all dimensions $d$ because of the localization phenomenon. However, Thouless and coworkers$^{3,4}$ argued that $G$ at any length scale should still be determined by $G$ at any one length scale, as long as this length scale is longer than the microscopic length scale (which is the elastic mean free path in the localization problem). To make this argument they had to formulate a definition of the conductance at any length scale. The definition proposed was that the conductance, in dimensionless units, of a system of volume $L^d$ is simply the ratio $g(L) = \Delta E(L)/\partial E/\partial N(L)$, where $\Delta E$ is the geometric mean of the shift in the system's energy levels due to a change in its boundary conditions, and $\partial E/\partial N$ is the average energy level spacing. In the limit of weak disorder, this definition can be shown to agree with Ohm's law using an argument based on the uncertainty principle$^5$, or, more rigorously, from the Kubo formula$^2$. In the strong disorder limit it has the appropriate qualitative features since if the states are localized, most of them should be exponentially insensitive to the boundary conditions. However, the power of this definition was that one could use it to argue that $g(L)$ should determine $g$ at any scale $bL$. Imagine joining together blocks of size $L^d$ to make blocks of size $(2L)^d$ and calculating $g(2L)$ by using perturbation theory, taking the isolated blocks as your unperturbed system. The new value $g(2L)$ will be determined by (1) the proximity in energy of the levels of the unperturbed systems, i.e. $\partial E/\partial N(L)$, and (2) the size of the matrix elements connecting the states across the boundaries, i.e.
\[ \Delta E(L) \]. And, although there existed no rigorous proof, it was not hard to believe that in fact it was only the ratio of these two quantities, \( g(L) \), which mattered.

The full consequences of this scaling hypothesis were not exploited until very recently, when the notorious "Gang of Four" used it to argue that no Anderson transition should occur for \( d < 2 \), and that all states should be localized for any degree of disorder in these dimensionalities.\(^6\) The argument they gave was deceptively simple. The scaling hypothesis says that \( g(bL) = f[b, g(L)] \), or for an infinitesimal rescaling \( b \), that the logarithmic derivative of \( g \) is a function only of \( g \). Since \( d \ln g / d \ln L = (1/g) d g / d \ln L \), we can define

\[ \beta(g) \equiv \frac{d \ln g}{d \ln L} . \]  \tag{1-2} \]

We know \( \beta(g) \) in the two limits, \( g \to 0 \), and \( g \to \infty \). As \( g \to \infty \) (weak disorder), Ohm's law holds and from (1) and (2)

\[ \beta(g \to \infty) = (d - 2) . \]  \tag{1-3} \]

As \( g \to 0 \) (strong localization), \( g \propto e^{-L} \), and

\[ \beta(g \to 0) = \ln g . \]  \tag{1-4} \]

Moreover, we expect \( \beta(g) \) to be monotonic in \( g \) since if we consider rescaling several systems of the same initial size \( L^d \), but different \( g(L) \), we certainly expect the greatest positive change in \( g \) to occur in the system which couples most strongly at the boundaries. Furthermore,
since the rescaling transformation consists of fitting together finite blocks it cannot produce any singularities in $\beta(g)$ or its derivatives for finite $L$. Thus, $\beta(g)$ is continuous and monotonic; but that's all we need to know! Equations (1-3) and (1-4) then tell us that for $d \leq 2$, $\beta(g)$ is negative for all finite $g$. The sign of $\beta(g)$ determines whether the conductance increases or decreases as $L$ increases, so if $\beta(g) < 0$ for $d \leq 2$, then every system will scale to zero conductance as $L \to \infty$. We interpret this to mean that in the infinite system, all states are localized for $d \leq 2$. For $d > 2$, $\beta(g)$ is positive as $g \to \infty$, and an Anderson transition can occur at finite $g$. For a finite system with relatively weak disorder and $d \leq 2$, this universal scaling of $g$ to zero will manifest itself only as a small negative contribution to the Ohmic conductivity, which increases with the system size.

The experimental consequences of the scaling hypothesis at finite temperature are quite interesting and will be discussed in detail for the one-dimensional case in Chapters VII and VIII from a theoretical point of view. However, there were still many questions to be answered. First, there was the problem that the quantity $g$ was only conjectured to be the conductance of the system, it was not derived from linear response theory except in the weak disorder limit. Secondly, all the quantitative calculations of localization effects (and "interaction" effects, as we will see later), except in one dimension, were based on the impurity-averaged perturbation theory which we will describe and utilize in Chapter VIII. This approach is only valid in the weak disorder limit ($\epsilon_f \tau \gg 1$, where $\epsilon_f$ is the fermi energy and $\tau$ the elastic scattering time).
These problems, combined with the first successful experimental measurements of localization in thin metallic wires, has engendered renewed theoretical interest in the localization problem in one dimension. For the one-dimensional case, a number of quite old arguments existed which showed that all eigenstates should be localized if there were any disorder in the potential. Nonetheless, every ten years some heretics would arise and attempt to cast doubt on the universal validity of these arguments. In 1967, Theodorou and Cohen argued that there was one extended state in the 1d Anderson model. More recently, Czycholl and Kramer claimed to find numerical evidence for extended states in one-dimension. If one believes in the universality of the scaling hypothesis, one would certainly like to see if such cases exist in 1d and if they do, understand why they are exceptional. This is one of the questions addressed in Chapters III, IV and V.

More importantly, in one dimension, the weaknesses in the scaling argument mentioned above can be overcome. In one dimension a simple expression for the conductance at finite length scale, valid for any degree of disorder, can be derived from linear response theory. It states that the dimensionless DC conductance, g, of a disordered segment of length L is simply T/R where R and T are the total reflection and transmission coefficients of the segment. This relation was first derived by Landauer in 1970, and again by Anderson et al. in 1980; however, their arguments were not rigorous, and attempts to derive their result rigorously from linear response theory initially failed. Thouless then pointed out that these derivations failed to preserve
charge neutrality in the systems considered, and that a careful application of the Kubo formula gave Landauer's result.\textsuperscript{18} In Chapter II, we give a rigorous derivation of the Landauer formula, along the lines suggested by Thouless.

Using the Landauer formula, it is possible to derive an exact scaling law for \( g \) (or for some more suitable function of \( g \)) in various 1d disordered systems at zero temperature. Landauer,\textsuperscript{14} and later Anderson et al.\textsuperscript{15} calculated such scaling relations on the basis of a model of 1d scatterers with randomly phased scattering matrices. Although they obtained a number of appealing results, it was not clear how general these results were, since they did not begin from a random model Hamiltonian, as is usually done. And, in fact, the general validity of their random phase assumption has been challenged quite recently.\textsuperscript{19} In Chapter III, using the Landauer formula, we obtain an exact scaling law for the Anderson model Hamiltonian with both diagonal and off-diagonal disorder\textsuperscript{20} and show that it agrees with the random phase model in the limit of weak disorder. In Chapter IV, we explicitly derive the probability distributions for \( g \) and \( 1/g \) for the special case treated by Theodorou and Cohen.\textsuperscript{21} We show that, contrary to their claim, it does not correspond to an extended state. In Chapter V, we establish the equivalence of the Anderson model with weak disorder to the random-phase model by calculating the relevant phase distributions explicitly for the Anderson model, and argue that the random phase assumption is justified in the weak disorder limit for almost any model.\textsuperscript{22} In Chapter VI we discuss the unusual statistical properties of the distributions for \( g \) and
1/g which arise in 1d, and formulate techniques for calculating meaningful averages given such poorly behaved distributions. In Chapter VII we describe the effects of finite temperature in aiding conduction using ideas suggested by Thouless; and formulate an argument to show that the enormous intrinsic statistical fluctuations in the zero temperature response functions should not be important in the range of system parameters presently accessible experimentally. In Chapter VIII, we consider the results of experiments on low temperature conduction in thin metal wires and compare the predictions of the localization and "interaction" theories. Using impurity-averaged perturbation theory we calculate the inelastic scattering rate of diffusing electrons from acoustic phonons and from tunneling centers in disordered solids. In Chapter IX we compare the results with experiments, discuss the remaining discrepancies between theory and experiment, and make some concluding remarks about our present understanding of localization and low temperature transport in disordered 1d systems.

A final introductory remark about studying localization in 1d. It appears that the essential physics of localization is contained in the one-dimensional case. As an electron attempts to propagate through the random potential in one dimension, it suffers multiple reflections at each barrier it reaches. The random relative phases of these multiply-reflected waves causes them to cancel on the average. The only remaining contribution after N barriers is from the unscattered wave which decays exponentially as \( \exp[-N|\ln T|] \), where T is the transmission coefficient of a barrier, since the modulus of the unscattered wave is just \( |T|^N \).
This destructive interference of the back-scattered electron waves is the physical cause of the localization phenomena in any dimension, it is merely diluted in higher dimensions. 25
Chapter II. The Landauer Conductance Formula

In this chapter, we present a rigorous derivation of the Landauer formula from linear response theory which is valid for any non-interacting disordered system of electrons of length $L$. We also generalize the formula to the case of spin-flip scattering, which breaks time-reversal symmetry. Before launching into the minutiae of this derivation, however, let us review the argument of Landauer\textsuperscript{14} which gives the correct result for the one-channel case.

Landauer argued as follows: Consider electrons incident from the left of an array of obstacles of length $L$, with total reflection coefficient $R$. Averaging over several wavelengths, the density to the left is $1+R$ and that on the right is $1-R$, giving a density gradient across the region, $\frac{\partial n}{\partial x} = \frac{-2R}{L}$. The current across the region is the incident flux minus the reflected flux, $J = v(1-R)$, where $v$ is the electronic velocity. Using the diffusion equation, $J = -D \frac{\partial n}{\partial x}$, gives $D = \frac{vL}{2} \left( \frac{1-R}{R} \right)$; and using the Einstein relation between the conductivity and diffusion coefficient of a degenerate fermi gas we get

$$\sigma/L = \frac{e^2}{2m^*_F} \left( \frac{1-R}{R} \right). \quad (2-1)$$

Taking the left hand side as a definition of $G$, this is the correct one-channel Landauer formula, valid for a general 1d system. However, this argument itself is quite suspect, since it appeals to a diffusion equation—which is known to be rigorously correct only in the weak disorder limit. The argument of Anderson et al.\textsuperscript{15} does not make this assumption, but is no more rigorous, so the conditions for the
validity of (2-1) remained unclear.

This motivated attempts to derive (2-1) rigorously from linear response theory, and generalize it for many scattering channels. However, the initial results of these investigations were to find \( G = \frac{e^2}{2\pi h} T \) for one channel,\(^{16}\) and \( G = \frac{e^2}{2\pi h} \text{Tr} (t^t t) \) for the many-channel case,\(^{17}\) where \( t \) is the transmission matrix to be defined below. These formulae had the disturbing feature that in the limit of zero disorder, where \( T = 1 \), the resistance is not zero. Although it was initially argued that such a result could occur if one attempted to apply a field over a finite region of a perfect conductor, it will be clear from the derivation presented below that these answers arise from an unphysical definition of conductance. This was first pointed out in a comment by Thouless\(^{18}\) and a detailed many-channel generalization, which reduces to the correct answer in the one-channel case, was given by Langreth and Abrahams.\(^{26}\) However, their argument does not explicitly display much of the interesting physics involved in the correct derivation. The approach taken below, which does not appear in the literature, is based on Thouless' comment and the suggestions of Dr. Patrick A. Lee.

Consider a one-dimensional sample of length \( L_0 \) which is disordered only in the region \( 0 < x < L \), where \( L \ll L_0 \). Let us apply a harmonic field \( E(x)e^{-i\omega t} \) with arbitrary spatial dependence to the whole sample and imagine calculating \( \langle J(\omega, x) \rangle \), the exact quantum-mechanical expectation value of the induced current of non-interacting electrons, to linear order in \( E \). Since \( E \) is not uniform, there is no reason a priori to expect \( \langle J(\omega, x) \rangle \) to be independent of \( x \) and in general it will not be.
However, in such a situation, charge will build up over time in the sample creating large internal fields. But it is never physically sensible to ignore electronic interactions when a macroscopic charge build-up can occur. The physically sensible approach is to assume that whatever external field is applied, the system will create internal fields such that $\langle J(\omega) \rangle$ is independent of $x$. We then define the conductance of the disordered region by

$$G(\omega) \equiv \frac{\langle J(\omega) \rangle}{\mathcal{V}}, \quad (2-2a)$$

where

$$\mathcal{V} \equiv \int_{0}^{L} E(x) \, dx \quad , \quad (2-2b)$$

and $E(x)$ is the self-consistent field which generates a uniform $\langle J(\omega) \rangle$. To find $G$ we must calculate $\langle J(\omega,x) \rangle$ using linear response theory, solve for the $E(x)$ which makes $\langle J(\omega) \rangle$ uniform, and then use this $E$ to calculate $G$.

We are interested in $G$ in the limit $\omega \to 0$, however, the main subtlety in the derivation arises in defining this limit appropriately. We begin with a system of finite total length $L_0$, excited at finite frequency $\omega$. We must always require that $L_0$ be large enough so that $\omega$ is much larger than the energy level spacing between states which contribute to conduction. If this condition were violated, the conductance would depend greatly on details of the boundary conditions, and no meaningful physical parameter of the bulk system should be strongly sensitive to boundary conditions. The states which contribute to conduction are those near the fermi level and in $1d$ they have a typical spacing
$\Delta E \approx v_f / L_0$ (where $v_f$ is the fermi velocity and we have set $\hbar = 1$); thus we always require
\[ \frac{\omega L_0}{v_f} \gg 1 \quad (2-3) \]

On the other hand, since $L$, the length of the disordered region, is much less than $L_0$, we can still make $\omega$ small enough so that
\[ \frac{\omega L}{v_f} \ll 1 \quad (2-4) \]

We shall see that this condition is crucial, since it implies that the variation in the current is negligible across the disordered region. Thus the correct procedure is to calculate $\langle J(\omega, x) \rangle$ for finite $\omega$ and find the $E$ which makes $\langle J(\omega) \rangle$ uniform for $\frac{v_f}{L_0} \ll \omega \ll \frac{v_f}{L}$, not for $\omega = 0$.

We begin by choosing a gauge where $\phi = 0$, and switch the electric field on adiabatically so
\[ A(x, t) = \frac{1}{\omega + i\delta} E(x) e^{-i\omega t + \delta t} \quad (2-5) \]

and
\[ H = \frac{(P - eA)^2}{2m} + U(x) \quad (2-6) \]

where $U(x)$ is random for $0 \leq x \leq L$, and zero otherwise. The perturbation which is linear in $A$ is
\[ H' = -\frac{e}{2m} [PA + AP] e^{-i\omega t + \delta t} \quad (2-7) \]

and we wish to calculate
to linear order in $A$, where $\rho$ is the density matrix of the system. We assume $\rho$ can be expanded as

$$\rho = \rho_0 + \rho_0 e^{-i\omega t + \delta t},$$

where $\rho_0$ is the density matrix of the unperturbed system. This is of the form

$$\rho_0 = \sum_{\alpha} |\alpha\rangle\langle \alpha| f(\epsilon_{\alpha})$$

where $|\alpha\rangle$ are the exact eigenstates of the unperturbed system, $H_0 = H(A=0)$, and $f(\epsilon_{\alpha})$ is the Fermi-Dirac distribution function. We first consider only $T=0$, so $f(\epsilon_{\alpha}) = \theta(\epsilon_f - \epsilon_{\alpha})$. The linearized equation of motion for $\rho'(t)$ is

$$i\hbar \rho'/\partial t = [H_0, \rho'(t)] + [H'(t), \rho_0].$$

We use (2-9), cancel the harmonic factors on both sides of (2-11), and then take matrix elements of (2-11) in the basis of $H_0$ to find

$$\rho_{\alpha\beta}' = \frac{H'_{\alpha\beta} [\theta(\epsilon_f - \epsilon_{\beta'}) - \theta(\epsilon_f - \epsilon_{\alpha})]}{\omega_{\beta\alpha} - \omega - i\delta},$$

where $\omega_{\beta\alpha} = \epsilon_{\beta} - \epsilon_{\alpha}$. The electron current operator, including spin is

$$<x'|J(x)|x'> = -e<x'|n(x)v_{op} + v_{op} n(x)|x'>$$

$$= -\frac{e}{2} I[\delta(x-x')v_{op}(x) + v_{op}(x)\delta(x-x')],$$

where $I$ is the $2 \times 2$ identity matrix in spin space, and the electronic
velocity operator, in the presence of A, is \( \mathbf{v}_{\text{op}} = \frac{\mathbf{p} - \mathbf{eA}}{\mathbf{m}} \) (setting \( c = 1 \)). However, the term in \( \mathbf{v}_{\text{op}} \) proportional to A gives a purely reactive contribution to the current, and, although we ultimately need to calculate both the real and imaginary parts of \( \langle J \rangle \), we can simplify the calculation greatly by calculating the real part of \( \langle J \rangle \) first and then using the Kramers-Kronig relation to obtain its imaginary part. This allows us to take \( \mathbf{v}_{\text{op}}(x) = -i \frac{d}{dx} \) throughout the calculation.

Using (2-12) and (2-13) we calculate \( \langle J(\omega, x) \rangle = \text{Tr}(\rho J) \)

\[
\langle J(\omega, x) \rangle = \text{Tr}(\rho J) = \sum_{\alpha\beta} \left[ \frac{\theta(\epsilon_f - \epsilon_\beta) - \theta(\epsilon_f - \epsilon_\alpha)}{\omega_\beta - \omega - i\delta} \right] H_{\alpha\beta} J_{\beta\alpha}(x) \tag{2-14}
\]

\[
= \frac{-i e^2}{(2m)^2} \sum_{\alpha\beta} \left[ \frac{\theta(\epsilon_f - \epsilon_\beta) - \theta(\epsilon_f - \epsilon_\alpha)}{\omega_\beta - \omega - i\delta} \right] W_{\alpha\beta}(x) \int_{-L_0/2}^{L_0/2} dx' W_{\alpha\beta}(x') E(x') ,
\]

where \( W_{\alpha\beta}(x) = \psi_\alpha^*(x) \cdot \frac{\partial \psi_\beta}{\partial x} - \psi_\beta^*(x) \cdot \frac{\partial \psi_\alpha}{\partial x}(x) \), (the dot denotes the inner product in spin space). We have used (2-5), (2-7) and integrated by parts to obtain (2-14). Taking its real part and using some symmetries of \( W_{\alpha\beta} \), gives

\[
\langle J_r(\omega, x) \rangle = \frac{-ie^2}{(2m)^2} \sum_{\epsilon_f} \sum_{\epsilon_f' > \epsilon_f} \delta(\omega_\beta - \omega) W_{\alpha\beta}(x) \int_{-L_0/2}^{L_0/2} dx' W_{\alpha\beta}^*(x') E(x') \tag{2-15}
\]

\[
\equiv \int_{-L_0/2}^{L_0/2} dx' \sigma_r(x,x') E(x') ,
\]

where we have assumed \( \omega > 0 \) which eliminates one of the \( \delta \) functions. To calculate \( \sigma_r(x,x') \) we need to get an explicit expression for \( W_{\alpha\beta}(x) \), which appears to require knowing the exact eigenstates of the system, even for
x in the disordered region where $\psi_\alpha(x)$ is some horrible fluctuating beast. However, we shall see that current conservation saves the day by allowing us to show that despite the wild fluctuations in $\psi_\alpha(x)$ and $\psi_\beta(x)$, the value of $W_{\alpha\beta}(x)$ is essentially unchanged in the disordered region from its value in the ordered region. In the ordered region, the eigenstates are just the sum of plane waves incident on, and scattered from the disordered region. We can characterize the disordered region by an S-matrix of the form

$$S = \begin{pmatrix} r & \bar{t} \\ \bar{t} & r \end{pmatrix},$$

where $r, t$ and $\bar{r}, \bar{t}$ are $2 \times 2$ matrices (for the case of spin channels only), which give the reflected and transmitted fluxes for waves incident from the left and right respectively. Conservation of flux always implies that $S$ is unitary, which gives a variety of useful algebraic relations between the four transmission and reflection matrices. For a time reversal invariant Hamiltonian additional symmetries of $S$ exist which we shall not assume exist in completing our derivation. If we choose the eigenstates in the ordered region to correspond to a single electron with spin $|s>$ incident from either the right or left, they can be written in terms of the scattering matrix in the form

$$\psi^L_{ks}(x) = \begin{cases} \begin{align*} |s> e^{ikx/\sqrt{L_0}} + r|s> e^{-ikx/\sqrt{L_0}} & \quad x < 0 \\ t|s> e^{ikx/\sqrt{L_0}} & \quad x > L \end{align*} \end{cases}$$

(2-17a)
Thus a single index $\alpha$ in (2-15) corresponds to choosing an incident spin $|s\rangle$, wavevector $k > 0$, and a directional index, $d = L, R$. The energy $\varepsilon_\alpha = k^2/2m$, is independent of $s$ and $d$; so before performing the sum over $k, k'$ in (2-15), we would like to calculate the quantity

$$Q(k, k', x, x') = \sum_{d, d' = L, R} \sum_{s, s'} W^{dd'}_{kk', ss'}(x) W^{dd'*}_{kk', ss'}(x') . \quad (2-18)$$

Actually, we would like someone else to calculate it, since it is lengthy and rather tedious. You're in luck, I'm not. Substituting (2-17) into the definition of $W_{\alpha\beta}(x)$ and using only properties of $r$ and $t$ which follow from unitarity, we find

$$Q(k, k', x, x') = \frac{2(k + k')^2}{L_0} \left\{ 2 \cos [(k' - k)|x - x'|] \right.$$ 

$$- \text{Tr}(t^t r) \cos[(k' - k)(|x| + |x'|)] \right\} , \quad (2-19)$$

where the trace is over spin degrees of freedom, and we have neglected terms of order $\omega/\varepsilon_f$. Now, using (2-19) in (2-15), the only remaining sum is over $k' > k_f$ and $k < k_f$. Making the replacement $\sum_{k'} \rightarrow \frac{L_0}{2\pi} \int dk'$, we integrate over $k'$, eliminating the $\delta$-function to give
\[ \sigma_r(x,x') = \frac{e^2 k_f}{2\pi \omega} \int \frac{dk}{k_f(1 - \frac{\omega}{2e_f^2})} \text{Tr} \left\{ \cos \left( \frac{\omega}{v_f} |x-x'| \right) - r^+ r \cos \left( \frac{\omega}{v_f} (|x| + |x'|) \right) \right\} \]

\[ = \frac{e^2}{2\pi} \text{Tr} \left\{ \cos \left( \frac{\omega}{v_f} |x-x'| \right) - r^+ r \cos \left( \frac{\omega}{v_f} (|x| + |x'|) \right) \right\} \] \hspace{1cm} (2-20)

Since \( \sigma_r(x,x') \) is just a sum of cosines, the Kramers-Kronig relation to find the imaginary part of \( \sigma \) is trivial; the analytic function of \( z \) whose real part is \( \cos(z) \) is just \( e^{iz} \) for \( \text{Re}(z) > 0 \). Thus

\[ \sigma(x,x') = \frac{e^2}{2\pi} \text{Tr} \left\{ \exp \left( \frac{i\omega}{v_f} |x-x'| \right) - r^+ r \exp \left[ \frac{i\omega}{v_f} (|x| + |x'|) \right] \right\} \] \hspace{1cm} (2-21)

Note, however, that we are not allowed to take \( \omega \to 0 \) in (2-21), since \( |x| \) or \( |x'| \) can be of order \( L_0 \), and the important restriction \( \frac{\omega L_0}{v_f} > 1 \) still applies. Strictly speaking, (2-21) only holds for \( x,x' \) in the ordered region; however, we now show that current conservation allows us to use it even in the disordered region.

Consider the one-channel case for simplicity. Since \( \psi_\alpha \) and \( \psi_\beta \) satisfy the Schrodinger equation everywhere, by simple substitution we see that

\[ i \frac{\partial}{\partial t} \psi_\alpha^* (x,t) \psi_\beta (x,t) = -\frac{3}{\partial x} \left( \frac{1}{2m} W_{\alpha\beta}(x,t) \right) \] \hspace{1cm} (2-22)

\[ \psi_\alpha (x,t) = \psi_\alpha (x) e^{-i\epsilon_\alpha t}, \text{thus} \quad i \frac{\partial}{\partial t} (\psi_\alpha^* \psi_\beta) = (\epsilon_\beta - \epsilon_\alpha) \psi_\alpha^* (x) \psi_\beta (x), \text{and for states which contribute to conduction } \epsilon_\beta - \epsilon_\alpha = \omega. \] If the states \( \psi_\alpha \) and \( \psi_\beta \)
decay on the average with localization length $\lambda < L$ in the disordered region, then integrating (2-22) gives

$$\omega \lambda e^{-L/\lambda} = \frac{1}{m} (W_{\alpha\beta}(L) - W_{\alpha\beta}(0)).$$

Dividing by $W_{\alpha\beta}(0) \approx k_f |t|^2 \approx k_f e^{-L/\lambda}$ to find the fractional change in $W_{\alpha\beta}$ gives

$$\frac{W_{\alpha\beta}(L) - W_{\alpha\beta}(0)}{W_{\alpha\beta}(0)} \approx \frac{\omega \lambda}{V_f} << 1. \quad (2-23)$$

If $\lambda > L$, then $L$ replaces $\lambda$ in (2-23). So in either case the fractional change in $W_{\alpha\beta}$ between any two points in the disordered region can be made arbitrarily small, and we may assume (2-21) holds even for $x, x'$ in the disordered region.

Finally, we are ready to solve for the $E$ field which makes $\langle J \rangle$ uniform. Impose a uniform field $E_0 e^{-i\omega t}$ over the disordered region. Using (2-21) and (2-15), we find that

$$J_0(x, \omega) = \frac{e^2}{2\pi} E_0 L \text{Tr}(t^t t) e^{-i\omega |x|/v_f}. \quad (2-24)$$

We see that such a field creates a charge density wave of wavelength $2\pi v_f / \omega$, which is necessarily shorter than the system size, since $\omega L_0 / v_f >> 1$. If one were to ignore this restriction and set $\omega = 0$ in (2-24), one gets $G \propto \text{Tr}(t^t t)$; which is essentially the mistake made by Economou and Soukoulis, and Fisher and Lee. Since (2-24) represents a macroscopic charge fluctuation it cannot be ignored and we seek to find the "internal field" which cancels this fluctuation. Impose a second
field $E_1 e^{-|x|} e^{-i\omega t}$ over the ordered region where $\epsilon > 0$ at the end of the calculation; then integrating using (2-21) we find,

$$
<J_1(x,\omega)> = \frac{e^2}{2\pi} \left( \frac{2E_1 v_f}{i\omega} \right) \left[ 2 - Tr(t^+ t) e^{-i\omega/v_f|x|} \right]. \quad (2-25)
$$

If we choose $E_1 = \left\{ \frac{i\omega L}{2V_f} E_0 \right\} Tr(t^+ t)/Tr(r^+ r)$ and add (2-25) to (2-24) we find we have indeed succeeded in cancelling the charge density wave by imposing this additional field in the ordered region $90^\circ$ out of phase with the field in the disordered region. We conclude that if the field $E_0$ were externally imposed, the system would respond by creating the internal field $E_1$. The remaining current in steady state is, from (2-24) and (2-25)

$$
<J(\omega)> = <J_0(x,\omega)> + <J_1(x,\omega)> = \frac{e^2}{2\pi} E_0 L \left[ Tr(t^+ t)/Tr(r^+ r) \right]. \quad (2-26)
$$

And from the definitions (2-2) we have (restoring units)

$$
G = \frac{e^2}{2\pi} \frac{Tr(t^+ t)}{Tr(r^+ r)}, \quad (2-27)
$$

which, thankfully, reduces to the Landauer formula for the one channel case.

There are several points to notice about this derivation. First, (2-27) was derived without assuming a time-reversal invariant Hamiltonian in the disordered region. Thus, for the case of random spin-flip scattering it can describe a random 1d system in a different universality class from the one-channel, time-reversal invariant case.
Second, (2-27) can be generalized to finite temperature easily simply by using (2-10) for the finite temperature density matrix. The result (for one-channel) is

\[ G = \frac{e^2}{2\pi\hbar} \int_0^\infty dk \frac{2f}{ek} (e_k) \frac{T}{R}, \tag{2-28} \]

where T and R are the transmission and reflection coefficients at energy \( e_k \). Thus, essentially we just thermally average the \( T = 0 \) result. Although this formula is formally correct within the assumptions of our derivation, in fact it is of little use in treating a real finite temperature system. Equation (2-28) just includes thermal effects insofar as they change the occupancy of the non-interacting electronic states; it does not include the effect that thermal interactions have in changing the nature of those states. The true electronic states in the presence of thermal interactions can be extended (insofar as that notion can be defined in an interacting system), which is why, at very low temperatures, thermal effects aid conduction. This very important point of physics will be discussed in detail in later chapters.

Third, the simplicity of the result (2-27) appears to be due to the fact that the two spin channels have the same velocities;\(^ {28} \) if the incident electrons can have many longitudinal velocities there appears to be no simple expression for G in terms of \( t \) and \( r \) only. Unfortunately, the most sensible approximate model for a metallic wire of finite cross-section is to treat it as a system whose states are labelled by a discrete number of transverse quantum states with energy \( \epsilon_n \) and a longitudinal momentum \( k_n \) which satisfies \( \epsilon_n + \frac{k_n^2}{2m} = \epsilon_f \) for the states
contributing to conduction. Thus, each state has a different velocity $k_n/m$ and (2-27) does not apply. Therefore, there exists no convenient, rigorous, Landauer-type formula which describes a wire, and the use of expressions for $G$ based on "intuitive" generalizations of the Landauer formula to explain the behavior of thin metallic wires is not on entirely firm theoretical ground, even at $T = 0$. However, it is certain that the addition of many scattering channels will not change the qualitative behavior of $G$ at $T = 0$, that $G$ decreases exponentially with the length of the sample for any degree of disorder. We derive this behavior rigorously for the one-channel case in the following chapter.
Chapter III. Exact Scaling Law for Average Resistance of the ld
Anderson Model

In this chapter, we present a recursive averaging technique which allows us to calculate exactly the average resistance (not conductance) of the generalized ld Anderson model. The technique is based on a transfer matrix formalism which will be developed below. An approach to the ld transmission problem based on such a formalism was first taken by Erdos and Herndon. They obtained a number of numerical and analytic results, but could not at the time address the issues raised by the scaling theory of localization, nor did they appreciate the complexity of the statistical treatment necessary in ld.

The 1980 paper by Anderson et al. was the first to use the Landauer formula and address these issues. It is worth summarizing their approach since many of our subsequent discussions will refer to it. They use the one-channel Landauer formula (2-1) and define the dimensionless resistance (not resistivity) $\rho = R/T$. Then they consider a model ld system characterized only as a sequence of point scatterers described by S-matrices (with unspecified Hamiltonian). They assume that the effect of disorder in the Hamiltonian is to determine some microscopic length $\lambda_p$ such that the S-matrices of scatterers separated by distances greater than $\lambda_p$ have totally random relative phases. Moreover, they assume that this length $\lambda_p$ is shorter than the localization length $\lambda$, so that their analysis will be relevant in both the weakly and strongly localized regimes. Then, using the law for the total transmission amplitude of two scatterers in series, they can easily calculate the phase-averaged total resistance of two segments of the chain,
where $\rho_1$ and $\rho_2$ are the resistance of the component segments. This is a useful formula, for one immediately sees that if $\rho_1$ and $\rho_2$ are small, then the third term in (3-1) is negligible, and one gets linear growth of $\rho$ with the number of scatters $N$ (ohmic behavior); whereas when $\rho_1$ and $\rho_2$ are large, the third term dominates and $\rho$ grows exponentially with $N$. In fact, converting (3-1) into a differential equation and averaging now over the moduli of $\rho_1$, $\rho_2$ yields

\[ \langle \rho \rangle = \frac{1}{2} \left( e^{2\gamma L} - 1 \right) , \]  

(3-2)

where $\gamma$ is an unspecified parameter of the model which should be related to the ohmic resistance. Anderson et al. then argue that actually (3-2) is not a very good estimate of the typical resistance of the sample and the correct quantity to look at is

\[ \bar{\rho} = \exp[\langle \ln (1+\rho) \rangle] - 1 , \]  

(3-3)

where we note that for small $\rho$, $\bar{\rho}$ is the arithmetic mean, and for large $\rho$ it is the geometric mean. The exponent in (3-3) satisfies

\[ \langle \ln (1+\rho) \rangle = \gamma L , \]  

(3-4)

so $\bar{\rho}$ grows half as rapidly as $\langle \rho \rangle$ as $L \to \infty$. We shall have a great deal more to say about the interesting statistical properties of $\rho$ below. However, as noted earlier, even though the results of this model, which we shall refer to as the random phase model, are quite suggestive and interesting, they are not derived from a random Hamiltonian. This leaves open
the questions: 1) How general are the results of the model? 2) Is the uniform phase randomness condition satisfied given a model ld Hamiltonian and under what conditions? 3) What is the physical significance of the large intrinsic statistical fluctuations in the zero-temperature ld resistance. In this chapter, we begin to answer these questions by studying the generalized ld Anderson model of a disordered solid. A number of papers concerning the problem of deriving the statistical properties of \( \rho \) from a model random Hamiltonian appeared around the time this research was completed; and Abrahams and Stephen were the first to publish results in agreement with our own for the case of diagonal disorder. The results of this chapter, however, were obtained independent of any knowledge of this related work.

We consider the generalized one-dimensional Anderson model described by the Hamiltonian

\[
H = \sum_{n=-\infty}^{\infty} \varepsilon_n |n\rangle \langle n| + \sum_{n=-\infty}^{\infty} V_{n,n+1} (|n\rangle \langle n+1| + |n+1\rangle \langle n|)
\]

where \( \{|n\rangle\} \) form a tight-binding basis set. The site energies \( \{\varepsilon_n\} \) are assumed to be uncorrelated random variables distributed with an arbitrary probability density \( P(\varepsilon_n) \), symmetric around zero for \( 1 \leq n \leq N \), and to be identically zero outside this region. Similarly, \( \{V_{n,n+1}\} \) are taken to be uncorrelated random variables for \( 1 \leq n \leq N-1 \) and simply equal to a constant \( V_0 \) outside this region. Thus, the system consists of a disordered segment containing \( N \) atoms and of length \( (N-1)d \) (where \( d \) is the lattice spacing), embedded in an infinite, perfectly conducting, ordered chain. If we write the solutions of the time-independent
Schrödinger equation in the tight-binding basis, the eigenvalue equation for the wave-function amplitudes $a_n$ may be summarized in terms of a $2 \times 2$ matrix which we call the promotion matrix, as follows

$$
P(n) \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} E - \varepsilon_n & -V_n, n+1 \\ V_n, n+1 & E - \varepsilon_{n-1} \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} .
$$

(3-6)

So that the real space wave function amplitudes at each end of the disordered segment are related by

$$
\begin{pmatrix} a_{N+1} \\ a_N \end{pmatrix} = \left[ \prod_{i=1}^{\infty} \mathbf{P}(i) \right] \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = \mathbf{P}_N \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} .
$$

(3-7)

We now derive a useful expression for $\rho$, the dimensionless resistance of the disordered region, in terms of the matrix elements of $\mathbf{P}_N$ by relating $\mathbf{P}_N$ to the transfer matrix, $\mathbf{T}_N$.

The transfer matrix is defined so as to relate the amplitudes of the solutions of the Schrödinger equation in the asymptotic (ordered) regions on either side of the disordered segment.

These solutions are

$$
a_n = \alpha e^{ikdn} + \beta e^{-ikdn} \quad -\infty < n < 1
$$

(3-8a)

$$
a_n = \gamma e^{ikdn} + \delta e^{-ikdn} \quad n > N
$$

(3-8b)

with

$$
E = 2V_0 \cos kd
$$

(3-8c)

and we define $\mathbf{T}_N$ by
The scattering matrix, $S_N$, defined in Eq. (2-16), relates $A, D$, the incident amplitudes, to $B, C$, the outgoing amplitudes. Using this equation, we can relate $T_N$ to $S_N$. The simple result is

$$|T_{N12}^2| = \rho$$

(3-10)

so we have merely to relate $T_N$ to $P_N$ to get the desired result. Equations (3-8) give us a relation between $a_{N+1}, a_N$ and $C, D$ of the form

$$
\begin{pmatrix}
a_{N+1} \\
a_N
\end{pmatrix} =
\begin{pmatrix}
e^{-ikd} & e^{ikd} \\
e^{-ikNd} & 0
\end{pmatrix}
\begin{pmatrix}
a_N \\
C
\end{pmatrix} = \Lambda^{-1} \begin{pmatrix} D \\ C \end{pmatrix},
$$

(3-11)

and a similar relation for $a_1, a_0$ and $A, B$ with the phase matrix $\theta^{-1}$ simply replaced by the unit matrix. Substitution of these relations into (3-7) gives the result

$$T_N = \theta \Lambda^{-1} P_N \Lambda$$

(3-12)

where we note that $\Lambda^{-1} \neq \Lambda^\dagger$ in general. From (3-12) and (3-10) we find, using the relation $\det(P_N) = 1$, the general expression for the resistance of any one-dimensional Anderson model

$$\rho = \frac{1}{4-(E^2/V_0^2)} \left\{ (P_{11}^N)^2 + (P_{12}^N)^2 + (P_{21}^N)^2 + (P_{22}^N)^2 + \\
\frac{E}{V_0} (P_{11}^N - P_{22}^N)(P_{12}^N - P_{21}^N) \right\}.
$$

(3-13)

This expression is very convenient for performing both analytic and numerical calculations on the Anderson model which we proceed to do in the following chapters.
Averaging Eq. (3-13) gives an expression for the arithmetic mean, \( \langle \rho \rangle \), in terms of quantities \( \langle P^i_j \rangle_{N} \), which can be determined by a recursive method based on Eqs. (3-6) and (3-7). We first consider the average of Eq. (3-13) for arbitrary diagonal and off-diagonal disorder at \( E = 0 \). Equation (3-7) implies the recursion relations

\[
\begin{align*}
\frac{p_{11}^n}{n} &= \frac{(E-\varepsilon_n)}{V_{n,n+1}} \frac{p_{11}^{n-1}}{n-1} - \frac{V_{n-1,n}}{V_{n,n+1}} \frac{p_{11}^{n-2}}{n-2}, \\
\frac{p_{21}^n}{n} &= \frac{p_{11}^{n-1}}{n-1}, \\
\frac{p_{12}^n}{n} &= \frac{(E-\varepsilon_n)}{V_{n,n+1}} \frac{p_{12}^{n-1}}{n-1} - \frac{V_{n-1,n}}{V_{n,n+1}} \frac{p_{12}^{n-2}}{n-2}, \\
\frac{p_{22}^n}{n} &= \frac{p_{12}^{n-1}}{n-1}.
\end{align*}
\]

If we define

\[
F_n = \langle (p_{11}^n)^2 \rangle, \quad f_n = \langle (p_{12}^n)^2 \rangle,
\]

and set \( E = 0 \), then (3-13) takes the form

\[
\langle \rho \rangle = \frac{1}{4} \left[ F_N + F_{N-1} + f_N + f_{N-1} \right] - 1/2.
\]

From squaring Eq. (3-14) and averaging we obtain the same recursion relation for \( F_n \) and \( f_n \) which can be written in matrix form as

\[
\begin{pmatrix}
F_n \\
F_{n-1}
\end{pmatrix} = \begin{pmatrix}
\langle \varepsilon^2 \rangle & \langle \varepsilon \rangle \langle V \rangle \\
\langle V \rangle & \langle V^2 \rangle
\end{pmatrix}\begin{pmatrix}
F_{n-1} \\
F_{n-2}
\end{pmatrix}.
\]

(3-17)
where we have used $\langle \varepsilon \rangle = 0$ and the fact that $p_{n-1}^{11}$ depends only on $\varepsilon_{n-1} \cdots \varepsilon_1$, and $V_{n-1, n} \cdots V_{1,2}, V_0$. Denoting the 2x2 recursion matrix in (3-17) by $R$, Eq. (3-7) implies that

$$
\begin{pmatrix}
F_N \\
F_{N-1}
\end{pmatrix}
= R^{N-2}
\begin{pmatrix}
F_2 \\
F_1
\end{pmatrix}.
$$

(3-18)

Some care must be taken in the last iteration of the recursion relation since the promotion matrix at the boundary, $P^{(N)}$, depends on a non-random variable $V_{N, N+1} = V_0$ which is not to be averaged. The eigenvalues and eigenvectors of $R$ are found to be

$$
\lambda_{\pm} = \frac{1}{2} \langle \varepsilon^2 \rangle \frac{1}{V^2} \pm \left[ \left( \frac{\langle \varepsilon^2 \rangle}{2} \left( \frac{1}{V^2} \right) \right)^2 + \langle V^2 \rangle \left( \frac{1}{V^2} \right) \right]^{1/2}
$$

(3-19a)

$$
U_+ = \begin{pmatrix} 1 \\ 1/\lambda_+ \end{pmatrix}, \quad U_- = \begin{pmatrix} 1 \\ 1/\lambda_- \end{pmatrix}.
$$

(3-19b)

By expanding the initial vector of (3-17) as

$$
\begin{pmatrix}
F_2 \\
F_1
\end{pmatrix} = \alpha U_+ + \beta U_-,
$$

(3-20)

we see that

$$
\begin{pmatrix}
F_N \\
F_{N-1}
\end{pmatrix} = \alpha (\lambda_+)^{N-2} U_+ + \beta (\lambda_-)^{N-2} U_-.
$$

(3-21)

$f_N, f_{N-1}$ may be determined in an exactly analogous manner. Equations (3-19), (3-20), and (3-21) together with a careful treatment of the averages of promotion matrices at the boundaries (which contain non-random variables) yield the desired exact, general scaling law for $\langle \rho \rangle$, 

\[ \langle \rho \rangle = \frac{1}{4(\lambda_+ - \lambda_-)} \left\{ \left( 1 + \frac{\langle e^2 \rangle}{V^2} \right) \left( \lambda_+^{N} - \lambda_-^{N} \right) \right\} \]
\[ + \left( \frac{\langle V^2 \rangle}{V_0^2} + \frac{\langle e^2 \rangle}{V^2} \right) \left( \lambda_+^{N-1} - \lambda_-^{N-1} \right) \]
\[ + \langle V^2 \rangle \left( \frac{1}{V^2} \right) \left( \lambda_+^{N-2} - \lambda_-^{N-2} \right) \]
\[ - \frac{1}{2} \right\} . \]  

(3-22)

The salient features of this result are as follows: First, the eigenvalue \( \lambda_+ \) is always greater than or equal to the quantity \( \langle V^2 \rangle \langle \frac{1}{V^2} \rangle \), which in turn is greater than unity for any normalizable probability distribution of finite width by the Schwartz inequality. Since \( \langle e^2 \rangle > 0 \) for any degree of diagonal disorder, we see that \( \lambda_+ > 1 \) for any type or degree of disorder. Also \( \lambda_- \) is always negative and \( |\lambda_-| < \lambda_+ \). Therefore, in the limit \( N \to \infty \) Eq. (3-22) gives the result that in all cases the mean resistance grows exponentially with \( \langle \rho \rangle \sim \exp[N \ln \lambda_+] \). Second, \( \langle \rho \rangle \) depends only on the moments \( \langle e^2 \rangle \), \( \langle V^2 \rangle \) and \( \langle V^{-2} \rangle \) of \( P(e) \) and \( Q(V) \) so we see that our result is relatively insensitive to the precise nature of these distributions. Third, in the limit of weak disorder as \( N \to \infty \), Eq. (3-22) reduces to

\[ \langle \rho \rangle = \frac{1}{2} \left[ e^{N \ln \lambda_+} - 1 \right] , \]

(3-23)

which is the form of the result which Landauer\textsuperscript{14} and Anderson et al.\textsuperscript{15} derived for the random phase model. This suggests that the Anderson model and the random phase model may be equivalent in the limit of weak disorder. We shall demonstrate this equivalence in Chapter V; and discuss the generality and correctness of the assumptions underlying the random phase model in detail.
For purely diagonal disorder which has been considered by Abrahams
and Stephen\(^{34}\) \(\lambda_- < 1\) and so terms proportional to \((\lambda_-)^N\) decay exponentially. Dropping these terms and setting \(<\mathcal{V}^2> <\mathcal{V}^{-2}> = 1\) gives

\[
<\rho>_{\text{diag}} = \frac{1 + \left[\frac{<\mathcal{V}^2>}{2V_0} + 1\right]^{\frac{1}{2}}}{\left(4\left[\frac{<\mathcal{V}^2>}{2V_0} + 1\right]^{\frac{1}{2}}\right)^{\frac{1}{2}}} \exp\left\{N \ln\left[\frac{<\mathcal{V}^2>}{2V_0} + 1 + \left[\frac{<\mathcal{V}^2>}{2V_0}\right]^2\right]\right\} - \frac{1}{2}. \tag{3-24}
\]

For purely off-diagonal disorder \(-\lambda_- = \lambda_+ = \left[<\mathcal{V}^2> <\mathcal{V}^{-2}>\right]^{\frac{1}{2}}\) and (3-22) gives the result.

\[
<\rho> = \frac{1}{2} \left[\frac{<\mathcal{V}^2>}{<\mathcal{V}^2>} - 1\right]^{-1} \quad \text{N even} \tag{3-25a}
\]

\[
<\rho> = \frac{1}{4} \left[\left(\frac{<\mathcal{V}^2>}{<\mathcal{V}^2> - 1} + \frac{V_0^2}{<\mathcal{V}^2>} - 1\right) \left[\frac{<\mathcal{V}^2>}{<\mathcal{V}^2>} - 1\right]^{\frac{1}{2}}\right] - \frac{1}{2} \quad \text{N odd}. \tag{3-25b}
\]

The difference between chains with an even or odd number of atoms arises from different boundary scattering in the two cases. Again, we note that \(\ln<\rho> \propto N\) as \(N \to \infty\), even for purely off-diagonal disorder at \(E = 0\) where numerical results\(^{35}\) have indicated \(\ln<\rho> \propto N^{\frac{1}{2}}\). This disagreement is due to the fact that \(<\rho>\) is skewed by a few very large, very improbable values which will not show up at all in numerical averaging over an ensemble of reasonable size. The special case of purely off-diagonal disorder at \(E = 0\) will be solved exactly in the next chapter. The skewing of the average and its significance for numerical calculations will be discussed in detail in Chapter VI.
Next we consider the case $E \neq 0$ which requires treating Eq. (3-13) in its full generality. Besides generalizing the recursion relation for $\langle (P^i)^2 \rangle$ there are in principle five new quantities $\langle P_n^{ij} P_n^{jk} \rangle$ for which recursion relations may be derived. However, this treatment may be simplified in several ways. First, since $P^1_{n-1}$ satisfies the same recursion as $P^1_n$, and $P^1_{n-1} = P^1_n$, and $P^2_{n-1} = P^2_n$, the only difference between $\langle (P^1_n)^2 \rangle$ and $\langle (P^1_{n-1})^2 \rangle$, $\langle (P^2_{n+1})^2 \rangle$, $\langle (P^2_{n+1})^2 \rangle$ consists in different initial values $\langle (P^i_j)^2 \rangle$ which is a boundary effect. Since we will only concern ourselves with the rate of exponential growth of $\langle \rho \rangle$ we neglect this difference and simply consider $\langle (P^1_n)^2 \rangle$. Also, neglecting boundary effects $\langle P^1_n P^2_{n-1} \rangle = \langle P^1_n P^2_{n-1} \rangle$ and $\langle P^2_n P^2_{n-1} \rangle = \langle P^1_n P^2_{n-1} \rangle$. Finally, again since $P^1_n$ and $P^2_n$ satisfy the same recursion relations, neglecting boundary effects, $\langle P^1_n P^2_{n-1} \rangle$ satisfies the same recursion relations as $\langle (P^1_n)^2 \rangle$, and $\langle P^1_n P^2_{n+1} \rangle$ the same as $\langle P^1_n P^2_{n-1} \rangle = \langle P^1_n P^2_{n-1} \rangle$.

Thus, the only two quantities required to get the growth rate of $\langle \rho \rangle$, are $\langle (P^1_n)^2 \rangle$ and $\langle (P^1_n)^2 \rangle$ (so at this point we mercifully suppress the matrix indices).

\[
\langle (P_n^1)^2 \rangle = \left( \frac{(E-E_n)}{V_{n,n+1}} P_{n-1} + \frac{V_{n-1,n}}{V_{n,n+1}} P_{n-2} \right)^2
\]

\[
= (E^2 + \epsilon^2) \frac{1}{V^2} \langle (P_{n-1})^2 \rangle + \frac{1}{V^2} \langle (P_{n-2})^2 \rangle
\]

\[
- 2E \frac{1}{V^2} \langle V_{n-1,n} P_{n-1} P_{n-2} \rangle
\]

(3-26)

Since $P_{n-1}$ depends on $V_{n-1,n}$ we cannot immediately factor the last average, i.e. $V_{n-1,n} P_{n-2} \neq V_{n-1,n} P_{n-2}$. However, calculation of the
relevant recursion relations shows that \( \langle V_n P_{n-1} \rangle = \frac{1}{V} \langle P_{n-1} \rangle \), and
\( \langle P_{n-1} \rangle \) satisfies the recursion relation

\[
\langle P_{n-1} \rangle = E \frac{1}{\sqrt{V}} \langle (P_{n-1})^2 \rangle - \langle P_{n-1} P_{n-1} \rangle ,
\]

so we can write the recursion relations (3-26) and (3-27) in the form

\[
\begin{bmatrix}
\langle (P_n)^2 \rangle \\
\langle (P_{n-1})^2 \rangle \\
\langle P_{n-1} P_{n-1} \rangle
\end{bmatrix}
= 
\begin{bmatrix}
(E^2 + \langle \epsilon^2 \rangle) \frac{1}{\sqrt{V}} & \langle V^2 \rangle & \frac{1}{\sqrt{V}} - 2E \frac{1}{\sqrt{V}} & \frac{1}{2} & \langle (P_{n-1})^2 \rangle \\
1 & 0 & 0 & 0 & \langle (P_{n-2})^2 \rangle \\
E \frac{1}{\sqrt{V}} & 0 & -1 & 0 & \langle P_{n-1} P_{n-2} \rangle
\end{bmatrix}
\]

The eigenvalues of this recursion matrix are the roots of the cubic equation

\[
C(\lambda) = \left\{ \lambda^3 + \left[ 1 - (E^2 + \langle \epsilon^2 \rangle) \frac{1}{\sqrt{V^2}} \right] \lambda^2 + \right.
\]

\[
\left. + \left[ \frac{1}{\sqrt{V^2}} (E^2 - \langle \epsilon^2 \rangle - \langle V^2 \rangle) \right] \lambda - \langle V^2 \rangle \frac{1}{\sqrt{V^2}} \right\} = 0 .
\]

Since \( C(1) = 2 [1 - \langle V^2 \rangle \frac{1}{\sqrt{V^2}} - \epsilon^2 \frac{1}{\sqrt{V^2}}] < 0 \) and \( C(\lambda + \infty) > 0 \) there must always exist a positive root greater than unity, \( \lambda_+ \), which leads to exponential growth with \( \langle \rho \rangle \propto \exp[N \ln(\lambda_+)] \). A simple expression for \( \lambda_+ \) may be obtained in the limit of weak disorder but for arbitrary energy by linearizing Eq. (3-29) around \( \lambda = 1 \), which is the relevant root for zero disorder. We write \( \lambda_+ = 1 + \delta \lambda \), \( V = V_0 + \delta V \), and insert these expressions into (3-29) keeping lowest order terms in \( \delta \lambda \), \( \langle (\delta V)^2 \rangle \), and \( \langle \epsilon^2 \rangle \); solving for \( \delta \lambda \) gives
\begin{equation}
\lambda_+ = 1 + \delta \lambda = 1 + \frac{2}{4 - E^2/V_0^2} \left[ \frac{\langle \varepsilon^2 \rangle}{V_0^2} + \frac{4\langle (\delta V)^2 \rangle}{V_0^2} \right]
\end{equation}

(3-30)

where in this case we have assumed \( \langle \delta V \rangle = 0 \). Thus, expanding \( \ln[1 + \delta \lambda] \)

\begin{equation}
\langle \rho \rangle \propto \exp \left[ N \frac{2\langle \varepsilon^2 \rangle/V_0^2 + 4\langle \delta V^2 \rangle/V_0^2}{4 - E^2/V_0^2} \right].
\end{equation}

(3-31)

A more careful treatment of the case of weak diagonal disorder with \( E \neq 0 \) analogous to our exact treatment of the \( E = 0 \) case yields the result

\begin{equation}
\langle \rho(E) \rangle_{\text{diag}} = \frac{1}{2} \exp \left[ N \ln \left( 1 + \frac{2\langle \varepsilon^2 \rangle/V_0^2}{4 - E^2/V_0^2} \right) \right],
\end{equation}

(3-32)

where we have neglected terms of order unity.

We see that the resistance rises more rapidly as we approach \( E = 2V_0 \) reflecting the greater localization of the states near the band edge. At the band edge, \( R/T \) blows up since there exist no normalizable eigenstates with \( E > 2V_0 \) which have finite amplitudes in both ordered regions.

We have seen that examination of the arithmetic mean indicates the existence of localization with \( \ln \langle \rho \rangle \propto N \) in all cases, even for purely off–diagonal disorder at \( E = 0 \) (see Eq. (3-25)). However, for these systems as noted earlier, the arithmetic mean does not in general represent a "typical" value for the resistance. One indication of this behavior may be obtained by an examination of \( \langle \rho^2 \rangle \); if it grows more rapidly with \( N \) than \( \langle \rho \rangle^2 \) then the distribution becomes broader and broader relative to the mean as \( N \to \infty \). We can calculate \( \langle \rho^2 \rangle \) by the recursion method, and at \( E = 0 \) it is determined by the recursion relation.
The eigenvalues are the roots of the cubic equation

\[ 0 = \Gamma^3 - (\langle e^4 \rangle + 1) \Gamma^2 + \left[ \frac{1}{V^4} (\langle e^4 \rangle - V^4 - 6 \langle e^2 \rangle^2) \right] \Gamma + \langle V^4 \rangle \frac{1}{V^4} \]  

(3-34)

Solving for the largest root \( \lambda_+ \) in the limit of weak disorder yields

\[ \Gamma_+ = 1 + \sqrt{3} \langle e^2 \rangle + \frac{8 \langle (\delta V)^2 \rangle}{V^2} \]  

(3-35)

whereas, in the limit of weak disorder at \( E = 0 \) the largest eigenvalue of Eq. (3-19) satisfies

\[ (\lambda_+)^2 = \left[ \langle V^2 \rangle \frac{1}{V^2} \right]^2 + \frac{\langle e^2 \rangle}{2V^2} \frac{1}{V^2} = 1 + \frac{\langle e^2 \rangle}{V^2} + \frac{4 \langle (\delta V)^2 \rangle}{V^2} \]  

(3-36)

so \( \Gamma_+ > \lambda_+^2 \) and \( \langle \rho^2 \rangle/\langle \rho \rangle^2 = \exp[N \ln \left( \frac{\Gamma_+}{\lambda_+} \right)] \), which diverges as \( N \to \infty \). In addition, for the case of purely diagonal disorder, \( \Gamma \) is determined by the simpler equation

\[ \Gamma^3 - (1 + \langle e^2 \rangle/\langle V^4 \rangle) \Gamma^2 + \left( \langle e^4 \rangle/\langle V^4 \rangle - 1 - 6 \langle e^2 \rangle^2/\langle V^4 \rangle \right) \Gamma + 1 = 0 \]  

(3-37)

For a gaussian distribution \( P(e) \), \( \langle e^4 \rangle = 3 \langle e^2 \rangle \), and this cubic factors with negative one as a root. Then one can obtain a simple expression for \( \Gamma_+ \) for any degree of disorder.
\begin{equation}
\Gamma_+ = 1 + \frac{<e^4>}{2} + \left( <e^4> + \frac{<e^4>^2}{4} \right)^{1/2}.
\end{equation}

One can verify from Eqs. (3-19) and (3-39) that in this case \( \Gamma_+/\lambda_+^2 > 1 \) for any degree of disorder. Because of this poor behavior of the distribution of \( \rho \) as \( N \to \infty \) we shall see in Chapter VI that it is really the geometric and not the arithmetic mean of \( \rho \) which gives a value of \( \rho \) representative of the distribution. In addition, it is convenient to study not \( \langle \ln \rho \rangle \) but \( \langle \ln(1+\rho) \rangle \) which is non-negative and will have no pathologies associated with the perfectly conducting (\( \rho=0 \)) members of the ensemble.

It is not possible to calculate \( \langle \ln(1+\rho) \rangle \) for the Anderson model by the recursive method we have developed in this chapter. However, in Chapter V we will calculate this quantity in the limits of weak and strong disorder by a different method based on solving for the probability distributions of the phases of the elements of the transfer matrices which appear in Eqs. (3-10) and (3-11). In the next chapter, we solve completely for the probability distributions of \( \rho \) (and related quantities) in an interesting special case.
Chapter IV: Exact Solution for Off-Diagonal Disorder at Bandcenter

In this chapter we calculate the exact probability distribution of $\rho$, and several interesting functions of $\rho$, for the 1d Anderson model with purely off-diagonal disorder at $E = 0$. We already know from the recursive calculations of Chapter III that $<\rho>$ grows exponentially with $N$, the number of atoms in the chain, but the fractional variance, $\text{var}(\rho)/<\rho>^2$, diverges as $N \to \infty$. In this chapter, we show explicitly that the geometric mean $\bar{\rho} = \exp[\ln(1+\rho)]$ is in fact a better indicator of a typical value of $\rho$ than $<\rho>$ for this special case. We shall later argue that this is generally true; however, this special case is particularly interesting because $\bar{\rho}$ and $<\rho>$ depend differently on $N$. We find $\bar{\rho} \propto \exp[N^{1/2}]$, whereas $<\rho> \propto \exp[N]$. This unusual behavior of $\bar{\rho}$ causes the localization length (as usually defined) to diverge, which led Theodorou and Cohen to argue mistakenly that this state is extended.\textsuperscript{12} Licciardello and Fleishman\textsuperscript{36} subsequently argued that the wave-function for such a state should decay as $\exp[-\sqrt{x}/\lambda]$ and so the state was not truly extended; however, the argument they give is flawed.\textsuperscript{37} We show rigorously that the resistance is of order $\exp[N^{1/2}]$ with probability approaching unity as $N \to \infty$. Actually, despite the unusual behavior of $\bar{\rho}$, the distribution function of $\rho$ has many features characteristic of the general case; also, we shall see that the unusual behavior of $\bar{\rho}$ disappears immediately for $E \neq 0$. Thus, this state does not constitute an important counterexample to the universal scaling hypothesis.
We now consider a one-dimensional Anderson model described by the Hamiltonian
\[ H = \sum_{n=-\infty}^{\infty} V_{n,n+1} (|n><n+1| + |n+1><n|) \]  
(4-1)
where \{|n>\} are a tight-binding basis set and we have set the non-random site energies equal to zero. The hopping matrix elements, \(V_{n,n+1}\), are taken to be independent random variables with distribution \(Q(V)\) assumed to be symmetric around an unperturbed value \(V_0\) for \(1 \leq n \leq N\) and simply equal to \(V_0\) otherwise. The formalism described by Eqs. (3-2), (3-3) and (3-9) still applies, and we use it to calculate the probability distribution of \(\rho\). For the case of purely off-diagonal disorder at \(E=0\), the analysis of Eq. (3-13) simplifies greatly; the promotion matrices take the form
\[ p(n) = \begin{bmatrix} 0 & -V_{n-1,n} \\ V_{n,n+1} & 0 \end{bmatrix} \]  
(4-2)
So that for \(N\) even, we have from (3-7) and (3-13)
\[ \rho = \frac{1}{4} \left(x + \frac{1}{x'} - \frac{1}{2}\right), \]  
(4-3)
with
\[ x = \frac{(V_{12})^2 (V_{34})^2 \cdots (V_{N-1,N})^2}{(V_{23})^2 (V_{45})^2 \cdots (V_{N,N+1})^2} . \]  
(4-4)
It follows from (3-7) that \(x^{1/2} = |a_{N+1}|/|a_1|\), so following Ref. 12, we tentatively define the localization length by
Thus all quantities of interest may be obtained from the probability distribution for x, D(x). Before computing D(x) we note that \( \langle \rho \rangle \) may be obtained from (4-3) and (4-4) by inspection, and has the Landauer form (see Eq. (3-2))

\[
\langle \rho \rangle = \frac{1}{N^2} \ln \left( \frac{a_{N+1}}{a_1} \right) = \frac{1}{2N^2} \ln x .
\]  

(4-5)

This agrees with the result obtained by the recursive method (Eq. (3-25a)) where we have assumed N is odd and considered a chain of \( N+1 \) atoms.

We now proceed to calculate D(x) from Eq. (4-4). The hopping matrix elements \( V_{n,n+1} \) in a standard tight-binding model depend on an overlap integral which typically decays exponentially with the distance between adjacent sites. Randomness in the \( V_{n,n+1} \) will then arise from randomness in these interatomic distance, which can occur for example when impurities are randomly distributed on a wire. Thus, it usually more physically sensible to assume a distribution for these distances, i.e., for the \( \ln(V_{n,n+1}) \) than for \( V_{n,n+1} \) itself; we will assume a gaussian distribution

\[
F(\ln V) = (2\pi \sigma^2)^{-1/2} e^{-(\ln(V/V_0))^2/2\sigma^2} .
\]  

(4-7)

It follows from (4-7) and (4-4) that the distribution of \( \ln x \) is also gaussian but peaked around zero,
\( G(lnx) = \left[8\pi\sigma^2N\right]^{-1/2} e^{-\frac{(lnx)^2}{8\sigma^2}}, \) \hspace{1cm} (4-8)

and finally that
\[ D(x) = \left[8\pi\sigma^2N\right]^{-1/2} \frac{e^{-\frac{(lnx)^2}{8\sigma^2}}}{x} \quad x \leq 0. \] \hspace{1cm} (4-9)

It is important to realize that for large \( N \) this result is independent of the choice of \( F(lnV) \); any well-behaved choice of \( F \) which is symmetric around \( lnV_0 \) according to the central limit theorem implies that \( G(lnx) \) becomes gaussian around zero as \( N \to \infty \). The advantage of our choice of \( F \) is that (4-8) and (4-9) are exact for all \( N \). \( D(x) \) is a simple example of a lognormal distribution, for \( N \gg 1 \) it has some extraordinary properties. It goes to zero at the origin and at infinity faster than any power of \( x \), so all its moments are defined. It has a huge maximum at \( x = \exp[-4\sigma^2] \) but falls to \( \left[8\pi\sigma^2N\right]^{-1/2} \) at \( x = 1 \), and has an enormous tail decaying as \( 1/x \) until \( lnx = \sqrt{8\sigma^2} \). The \( \text{Prob}(x < 1) = \text{Prob}(x > 1) \), but the integrated weight near \( x = 1 \) is negligible and all the weight in \( D(x) \) is at very large or very small values of order \( \exp[\pm \sqrt{8\sigma^2}] \) (this is at first surprising since the maximum in \( D(x) \) lies outside these regions). Since Eq. (4-3) for \( \rho(x) \) is symmetric as \( x \to 1/x \) this means that \( \rho \) must have a high probability of being large, consistent with our calculation of \( <\rho> \). Using (4-3) and (4-9) we find the distribution \( H(\rho) \) to be
\[ H(\rho) = [32\pi\sigma^2 \rho(\rho+1)]^{-1/2} \exp \left\{ \frac{-[ln(1+2\rho+2\sqrt{\rho(\rho+1)})]^2}{8\sigma^2} \right\}. \] \hspace{1cm} (4-10)
This distribution is lognormal for large \( p \), but for small \( p \), \( H(p) = p^{-\frac{1}{2}} \exp[-p/2N\sigma^2] \), which has a square root singularity at \( p = 0 \). This singularity is a quite general consequence of the form of \( \rho(x) \) which has a minimum at \( x = 1 \). This causes a singularity in the probability density analogous to Van Hove singularities in the density of states. This means that the perfectly conducting (\( p = 0 \)) members of the ensemble are most probable. Nonetheless this is misleading since the singularity is integrable and, in fact (consistent with what we found earlier for \( D(x) \)) we can show that \( \text{Prob}\{\exp[N^{1/2-\epsilon}] < p < \exp[N^{1/2+\epsilon}]\} \to 1 \) for any \( \epsilon > 0 \) as \( N \to \infty \).

It is clear from (4-10) that the contribution to \( \int_0^{\infty} H(p)dp \) from \( p < 1 \) goes to zero, as \( N \to \infty \), thus we may use the approximate form \( H(p) \approx \frac{1}{2} D(4p) \). Defining \( y = N^{-1/2} \ln(4p) \)

\[
\int_0^{\exp[N^{1/2-\epsilon}] \frac{1}{2} D(4p) dp = (32\pi \sigma^2)^{-\frac{1}{2}} \int_0^{N^{-\epsilon}} e^{-y^2/8\sigma^2} dy \to 0 \quad \text{as} \quad N \to \infty \quad (4-11)
\]

and an exactly similar calculation shows \( \text{Prob}\{p > \exp[N^{1/2+\epsilon}]\} \to 0 \) as \( N \to \infty \). Thus although the long tail of \( H(p) \) skews the average of \( p \) to be of order \( e^N \), such large values are quite improbable and a "typical" value of \( p \) is of order \( \exp[\sqrt{N}] \). Averages of \( p \) and \( p^2 \) may be found from \( D(x) \) and \( \rho(x) \).

\[
<\rho> = \frac{1}{2} \left[ e^{2N\sigma^2} - 1 \right] \quad , \quad (4-12)
\]

\[
<\rho^2> = \frac{1}{8} \left\{ e^{8N\sigma^2} - 4e^{2N\sigma^2} + 3 \right\} \quad , \quad (4-13)
\]

which is consistent with Eq. (4-6). Since \( <\rho^2>/<\rho>^2 \approx \exp[4N\sigma^2] \), the
fractional variance diverges as \( N \to \infty \), in agreement with the poor behavior of \( H(\rho) \) found in other one-dimensional systems.\(^{15,30,34}\) The distribution of \( g = 1/\rho, I(g) = H(1/g)1/g^2 \), is even more poorly behaved. The square root singularity in \( H(\rho) \) at \( \rho = 0 \) is reflected in the very slow decay \( I(g) \propto g^{-3/2} \) as \( g \to \infty \); thus all moments of \( g \) diverge. Nonetheless, again this is a misleading indication that the conductance is likely to be large since \( \text{Prob}(g > \exp[-N^{3/2}]) \to 0 \) as \( N \to \infty \).

The moments of the distribution of \( \ln[1+\rho] \) are most easily computed from \( D(x) \) and \( \ln[1+\rho] = \ln[(1+x)^2/4x] \). The integrals involved are difficult to do exactly analytically but may be rigorously bounded above and below. We find

\[
\frac{\Gamma(n+1/2)^2}{\Gamma(n+1)} \left( \frac{8N\sigma^2}{\sqrt{\pi}} \right)^{n/2} > \langle \ln(1+\rho) \rangle^n > \frac{\Gamma(n+1/2)^2}{\Gamma(n+1)} \left( \frac{8N\sigma^2}{\sqrt{\pi}} \right)^{n/2} - O(N^{-2}). \tag{4-14}
\]

The bounds converge as \( N \to \infty \), implying that unlike the results of Anderson et al.\(^{15}\) for the random phase model, \( \langle \ln(1+\rho) \rangle \propto \frac{1}{N^{1/2}} \), not \( N \), and the fractional variance of \( \ln(1+\rho) \) goes to a constant, not zero, as \( N \to \infty \). Nonetheless, this is a sufficient condition for \( \langle \ln(1+\rho) \rangle \) to be representative of the distribution, and illustrates the correctness of the arguments of Anderson et al.\(^{15}\) that the scale resistance, \( \bar{\rho} \), is the appropriate variable for characterizing the poorly behaved distribution \( H(\rho) \). The general argument is that one finds a variable \( f(\rho) \) which satisfies \( \langle f(\rho) \rangle \propto N \) and whose fractional variance is well-behaved as \( N \to \infty \). Then, the scale resistance, \( \bar{\rho} \), defined by \( \bar{\rho} = f^{-1}[\langle f(\rho) \rangle] \), will be representative of \( H(\rho) \). For this model \( f(\rho) \) turns out to be \( [\ln(1+\rho)]^2 \), which satisfies \( \langle [\ln(1+\rho)]^2 \rangle = 4\sigma^2 N \), so we define
\[ \bar{\sigma} = \exp\left[\left(\langle\ln(1+\rho)\rangle^2\right)^{1/2}\right] - 1 = \exp[2\sigma \sqrt{N}] - 1 \quad . \tag{4-15} \]

This new scaling law is entirely consistent with the result for a typical value of \( \rho \) we obtained from an analysis of the entire distribution, \( H(\rho) \), that the resistance is of order \( \exp[\sqrt{N}] \) with very high probability. On the other hand, (4-8) implies that the inverse localization length, as defined in (4-5), has a gaussian distribution peaked around zero, with vanishing width as \( N \to \infty \), so that \( \xi \to \infty \) with probability unity. We conclude that this definition of the localization length is a misleading indicator of the transport properties of this system. In general, if there exists an \( f(\rho) \) satisfying the two conditions stated above, then the localization length is more appropriately defined as

\[ 1/\bar{\xi} = \frac{1}{L} f(\bar{\sigma}) \quad , \tag{4-16} \]

which gives \( \bar{\xi} = d/4\sigma^2 \) in this case.

Putting these amusing subtleties side, the most important features of the distribution of \( \rho \) is that it is lognormal in the tail, and \( \exp[\langle\ln(1+\rho)\rangle] \) gives a value typical of the distribution. In the general case, for reasons we will discuss below, we expect the distribution of \( \ln(1+\rho) \) to be approximately normal everywhere, and \( \exp[\langle\ln(1+\rho)\rangle] \) to be typical of the distribution, but \( \langle\ln(1+\rho)\rangle \propto N \), not \( N^{1/2} \). We first show numerically that the behavior \( \langle\ln(1+\rho)\rangle \propto N^{1/2} \) disappears as soon as we take \( E \neq 0 \), or if we introduce any diagonal disorder along with the off-diagonal disorder at \( E = 0 \).

Our calculations are again based on Eq. (3-13) for \( \rho \) in terms of the matrix elements of the promotion matrix. We computed numerically
the promotion matrices of an ensemble of chains of varying lengths $N$, and then calculated the quantities $\overline{\ln(l+p)}$ and $\ln[l+p]$, where the bar bracket denotes the ensemble average. Off-diagonal disorder was introduced into the promotion matrices by randomly generating the hopping matrix elements $V_{n,n+1}$ according to a rectangular probability distribution of width $2W$ centered around $V_0$ which we set equal to unity. Our basic result is that $\ln(l+p) = N$ as $N \to \infty$ for all $E \neq 0$ even down to $E = 10^{-4}V_0$, so that the behavior of the geometric mean for off-diagonal disorder at $E = 0$ is a very special case, at least in one dimension.

In Fig. 1 we present a plot of $\ln(l+p)$ versus $N$ with $W_V = .25$ for several non-zero values of $E$. Note that for all non-zero energies even for $E = .01$ we get a very precise linear dependence down to $N = 100$.

Qualitatively, this kind of behavior for $E \neq 0$ is not too surprising, as may be seen by a consideration of the recursion relations for $P_{n1}$. For diagonal disorder at $E = 0$ the recursion relation is

$$P_{n1} = -\frac{\varepsilon_n}{V_0} P_{n-1}^{\text{II}} - P_{n-2}^{\text{II}}$$

which has already been shown numerically to lead to the behavior $<\ln(l+p)> \propto N^{1/2}$. For off-diagonal disorder at $E \neq 0$ the recursion relation is

$$P_{n}^{\text{II}} = \frac{E}{V_{n,n+1}} P_{n-1}^{\text{II}} - \frac{V_{n-1,n}}{V_{n,n+1}} P_{n-2}^{\text{II}}.$$  

If one considers $E = V_0 = 1$ and takes a rectangular distribution of $\varepsilon_n$ of width $W_{E} = 1$ centered at zero and a similar distribution of $V_{n,n+1}$ centered at one then trivially the random variable $E/V_{n,n+1}$ is always positive and greater than $\varepsilon_n/V_0$; thus one expects $P_{n}^{\text{II}}$ for off-diagonal to grow faster than for diagonal disorder for equal degrees of disorder and $E$ of order unity. It does not seem obvious however, that for $E << V_0$ we should get such precise linear dependence, and it would appear that more
powerful mathematical techniques such as those employed by O'Connor are necessary to show this, as will be discussed in the next chapter.

The slope of the lines in Fig. 1 is the inverse localization length \( \frac{1}{N \ln(1+p)} \), and it is a monotonically increasing function of energy as expected since the states should be more strongly localized nearer the band edge. The results are consistent with the energy dependence

\[
\frac{1}{N} \ln(1+p) \approx \frac{1}{4-E^2};
\]

which is the analytically derived energy dependence of \( \frac{1}{N} \ln[1+p] \) for weak disorder, and the correct energy dependence for weak diagonal disorder, as we show in Chapter V.

At \( E = 0 \) we do get the result \( \ln(1+p) \approx \sqrt{N} \) in good agreement with our analytic result (Eq. 4-14) and the numerical results of Ref. 35. These results are shown in Fig. 2 (the solid line is the analytic result). The analytic result is only exact as \( N \to \infty \), and we found that for weak disorder (\( W \geq .05 \)) we had to go to chains as long as 10,000 atoms before a clear square root dependence was found. The behavior \( \ln(1+p) \approx \sqrt{N} \) is reached much more rapidly for large disorder (\( W \geq .5 \)), but then the onset of linear dependence for finite \( E \) also occurs more quickly. We note also that very large off-diagonal disorder (\( W > 1 \)) is somewhat unrealistic physically since in such a situation the chains have a large probability density near zero coupling which for example, causes the arithmetic mean of \( \rho \) to diverge. It is significant that there is substantially greater scatter in the data at \( E = 0 \) than for \( E \neq 0 \). We showed earlier that for off-diagonal disorder at \( E = 0 \) the relative variance of \( \ln(1+p) \), which is the ratio of the variance to the square of the mean, was independent of \( N \), unlike the case for a
gaussian distribution with non-zero mean, where this ratio decreases as \( N^{-1} \). We computed this ratio, \( \Gamma \), given by \( \Gamma = \frac{(\ln(1+p) - \ln(1+p))}{\ln(1+p)} \) for \( E \neq 0 \) and found that it did in fact decrease as \( N^{-1} \) as \( N \to \infty \) (see insert to Fig. 1), whereas for \( E = 0 \) it converged to the analytic result \( \frac{\pi}{2} - 1 \) and remained constant as \( N \) increased. This behavior was reflected in the variation of \( \ln(1+p) \) from ensemble to ensemble; we found that for ensembles of 500 chains of length 3200, the value of \( \ln(1+p) \) typically varied by one \( \pm 0.5\% \) from ensemble to ensemble for \( E = 0.5 \), whereas for \( E = 0 \) it varied by \( \pm 5\% \).

The behavior of the \( \ln(1+p) \) was much worse as expected. For an ensemble of 1000 chains of length 400 at \( E = 0 \) and \( W_v = 0.25 \) the value of \( \ln(1+p) \) typically varied by \( \pm 15\% \) as compared to \( \pm 2\% \) for \( \ln(1+p) \). The growth of \( \ln(1+p) \) was approximately linear with \( N \) and in fair agreement with the analytic result Eq. (4-12) up until \( N = 400 \) at which point it begins to grow more slowly roughly as \( \sqrt{N} \) with large fluctuations. The explanation for this behavior is that as \( N \) increases the numerical averaging procedure is seeing less of the long tail of the distribution, as will be discussed in detail in Chapter VI.

We next consider \( \ln(1+p) \) numerically for the Anderson model with both off-diagonal and diagonal disorder at \( E = 0 \). Again, previous work on the Anderson model with only diagonal disorder suggested that we should find \( \ln(1+p) \propto N \) for large diagonal disorder, and similar to our previous results, we find that even a small amount of diagonal disorder gives a linear dependence on \( N \). In Fig. 3 we plot \( \ln(1+p) \) vs. \( N \) for \( W_v = 0.25 \) and various values of \( W_c \); note the linear dependence down to
$W_\varepsilon = 0.01$. We found that as we increased $W_V$ the linear behavior could be seen clearly for even small values of $W_\varepsilon$. The inverse localization length $\frac{1}{N} \ln(l+\rho)$ grew more rapidly with $W_\varepsilon$ than it did with energy in the case of purely off-diagonal disorder as one might expect (until very near the band edge). We also calculated $\ln(l+\rho)$ with rather interesting results as is shown in Fig. 4-4. We found good agreement with the analytic result for $<\ln(l+\rho)>$ of Eq. (3-22) (the solid line) for $N$ up to 250, after which we find that the numerical average is consistently less than the analytic result with a percentage deviation which increases with $N$. This is exactly the behavior one would expect based on the assumption that $\ln\rho$ is normally distributed so that a numerical averaging procedure underweights the long tail of the distribution in $\rho$. This kind of behavior is characteristic of one-dimensional disordered systems, for reasons which we discuss in detail in Chapter VI and has been found e.g. by Andereck and Abrahams$^{30}$ working on a different model. In Chapter VI we will analytically derive a scaling law for $\ln(l+\rho)$ which gives an excellent fit to the numerical results of Fig. 4-4. This law is based on the assumption that $\ln(l+\rho)$ is normally distributed. To show that this is true quite generally, we must first explore the generality of the random phase assumption of Anderson et al.$^{15}$; since if this assumption is correct, it is quite easy to show that $\ln(l+\rho)$ is normally distributed. This is the subject of the next chapter.
Chapter V. Validity of the Random Phase Hypothesis

In this chapter we explore the validity of the phase randomness assumption employed by Landauer, and then Anderson et al. in their analyses of 1d localization. This question is important for several reasons. First, one would like to know if the random phase model captures the essential physics of localization in one dimension (and possibly in higher dimensions as noted earlier, see Ref. 25). Second, the knowledge that the random phase assumption holds in a particular class of models will allow us to show in Chapter VI that \( \ln(1+p) \) is normally distributed for all these models. Third, knowledge of the probability distributions of the relevant phases of the S-matrix will allow us to calculate \( \langle \ln(1+p) \rangle \) for the Anderson model, something we could not do using the recursive method of Chapter III. And finally, we would like to settle a controversy which has arisen recently concerning precisely this question of the conditions for the validity of the phase randomness assumption.

Lambert and Thorpe have re-examined the phase randomness assumption in the context of a model of randomly spaced delta function scatterers of uniform strength (hereafter referred to as the \( \delta \)-function model). They find that in fact the relevant phases are not random over \( 2\pi \) for weak disorder, and that even for strong disorder, when they are random over \( 2\pi \), there is no length scale over which phase randomness is achieved, contrary to the suggestion of Anderson, et al. Lambert and Thorpe suggested that their conclusions hold equally well for the Anderson model. This claim appears to conflict with the results we obtained
in Chapter III for the Anderson model, that the arithmetic mean, \( <R/T> \), has exactly the same form as for the random phase model in the limit of weak disorder but not in the limit of strong disorder. All these considerations led us to directly examine the question of phase randomness in the Anderson model. We find that, as our earlier work suggested, the Anderson model with weak diagonal disorder is essentially equivalent to the random phase model for long chains. Moreover, we find that even though phase randomness is not present at each site, there is an independent length scale at which uniform phase randomness is achieved as conjectured by Anderson et al.\(^{15}\). Finally, in the limit of strong diagonal disorder we find that the relevant phases are never uniformly distributed. These results are the exact opposite of those obtained for the \( \delta \)-function model\(^{19}\) and these differences arise from the quite different way in which disorder is introduced into the two models.

This chapter is organized as follows: First we set up the general formalism for analyzing the influence of the relevant phase distribution, \( P(\nu) \), on the scaling properties of \( \ln[1+\rho] \) in the Anderson model. Next we study \( P(\nu) \) numerically in the limit of weak disorder and give a heuristic argument for the behavior of \( P(\nu) \). We exploit the uniformity of \( P(\nu) \) to calculate the mean and variance of the inverse localization length analytically by a new method. We then analyze \( P(\nu) \) in the limit of large disorder. In this limit we calculate \( P(\nu) \) both analytically and numerically and show that it is non-uniform and independent of \( N \). Finally, we compare our results with those obtained for the \( \delta \)-function model and give an explanation for the very different behavior of the two
models. We show that a model of \( \delta \)-function scatterers with random strengths will behave similarly to the Anderson model, and our results suggest that any \( ld \) model with a random potential of varying sign will behave like the random phase model in the limit of weak disorder.

We begin by presenting several useful general equations derived by Lambert and Thorpe.\(^{19}\) To do this we unfortunately must introduce some new notation. Consider an arbitrary disordered region of \( N \) scatterers embedded in an infinite ordered chain. As noted above, each scatterer can be described by a transfer matrix \( T^{(n)} \) which relates the momentum-space wave-function amplitudes on one side of the barrier to those on the other, and the entire disordered region can be described by a total transfer matrix \( T_N \), which satisfies

\[
T_N = \prod_{n=1}^{N} T^{(n)} = T^{(N)} T_{N-1} .
\]  

(5-1)

If we denote \( R/T \) for the whole chain by \( \rho_N \) and \( R/T \) for a single scatterer by \( Z_N \) then we can parameterize these matrices in the form\(^{40}\)

\[
T_N = \begin{pmatrix}
(1+\rho_N)^{\frac{1}{2}} e^{-i\alpha_N} & -i\beta_N \\
^{\frac{1}{2}} \rho_N e^{-i\phi_N} & (1+\rho_N)^{\frac{1}{2}} e^{i\alpha_N}
\end{pmatrix}
\]  

(5-2a)

\[
T^{(N)} = \begin{pmatrix}
(1+Z_N)^{\frac{1}{2}} e^{-i\theta_N} & -i\phi_N \\
^{\frac{1}{2}} Z_N e^{-i\phi_N} & (1+Z_N)^{\frac{1}{2}} e^{i\theta_N}
\end{pmatrix}
\]  

(5-2b)
Inserting Eqs. (5-2a) and (5-2b) into Eq. (5-1) yields the recursion relations

\[
\ln(1+\rho_N) = \ln(1+\rho_{N-1}) + \ln(1+Z_N) + \ln(1+t_N^{1/2} \cos \nu_N) , \quad (5-3a)
\]

\[
a_N e^{i(\nu_N+1/2)} = \frac{b_N + a_{N-1} e^{i\nu_N}}{1 + b_N a_{N-1} e^{i\nu_N}} , \quad (5-3b)
\]

where

\[
\nu_N = \alpha_{N-1} + \beta_{N-1} + \theta_N - \phi_N ,
\]

\[
\delta_N = \theta_N + \theta_{N-1} - \phi_N + \phi_{N-1} , \quad (5-4)
\]

\[
b_N = \left( \frac{Z_N}{1+Z_N} \right)^{1/2} , \quad a_N = \left( \frac{\rho_N}{1+\rho_N} \right)^{1/2} , \quad t_N^{1/2} = b_N a_{N-1} .
\]

These are equations (5a), (5b) and (6) of reference 19. Iterating the recursion relation (5-3a) yields the exact result

\[
\ln(1+\rho_N) = N\gamma = N(\gamma_1 + \gamma_2) , \quad (5-5)
\]

with

\[
\gamma_1 = \frac{1}{N} \sum_{n=1}^{N} \ln(1+Z_n) , \quad (5-6a)
\]

\[
\gamma_2 = \frac{1}{N} \sum_{n=2}^{N} \ln(1+t_n^{1/2} \cos \nu_n) . \quad (5-6b)
\]

These are equations (7), (8) and (9) of Ref. 19. We see that it is the phase \( \nu_n \), defined in Eq. (5-4), that determines the scaling properties
of $\gamma_2$, while $\gamma_1$ is phase-independent. If $v_n$ is uniformly distributed over $2\pi$ for all $n$ and independent of $t_n$, then $<\gamma_2>$ is exactly zero, as can be verified by integrating Eq. (5-6b). Then $<\gamma> = <\gamma_1>$ and Eqs. (5-5) and (5-6) generate all the results of the random phase model. Thus, we will be primarily interested in determining the probability distribution $P_N(v_N)$, for ensembles of chains of varying lengths $N$, described by the Anderson model.

We consider the one-dimensional Anderson model with diagonal disorder described by the Hamiltonian

$$H = \sum_{n=-\infty}^{\infty} \varepsilon_n |n><n| + \sum_{n=-\infty}^{\infty} V \left[ |n><n+1| + |n+1><n| \right]. \quad (5-7)$$

The site energies $\{\varepsilon_n\}$ are now assumed to be independent random variables distributed uniformly over a width $2W$ for $1 \leq n \leq N$, and zero otherwise.

To analyze Eq. (5-3) we need to know $T^{(N)}$, the transfer matrix across the $N^{th}$ site. In Chapter III we showed (Eqs. 3-12) that the total transfer matrix across all $N$ sites was given by

$$T_N = \omega_N^{-1} P_N \omega_N^{1}, \quad (5-8)$$

where $\omega_N$ is a diagonal unitary matrix whose phases depend on $N$ (we have changed notation slightly from Eq. (3-12) to avoid ambiguities). $P_N$ is, as usual, the product of the promotion matrices for the disordered segment. It follows from (5-1) and (5-8) that we must make the definition,

$$T^{(N)} = \omega_N^{-1} P^{(N)} \omega_N^{1} \omega_{N-1}^{-1}. \quad (5-9)$$
Thus, 

\[ T(N) = \begin{pmatrix} 
1 - \frac{i\epsilon_N}{(4-E^2)^{1/2}} & \frac{i\epsilon_N}{(4-E^2)^{1/2}} e^{i(2Nkd-\pi/2)} \\
\frac{\epsilon_N}{(4-E^2)^{1/2}} e^{-i(2Nkd-\pi/2)} & 1 + \frac{i\epsilon_N}{(4-E^2)^{1/2}} 
\end{pmatrix} \]  

(5-10)

Comparison of Eq. (5-10) with Eq. (5-2b) gives the necessary relations

\[ \theta_N = \tan^{-1} \left[ \frac{\epsilon_N}{(4-E^2)^{1/2}} \right] , \]  

(5-11a)

\[ \phi_N = \pi/2 - 2Nkd + \pi/2 \left( 1 - \frac{\epsilon_N}{|\epsilon_N|} \right) , \]  

(5-11b)

\[ Z_N = \frac{\epsilon_N^2}{(4-E^2)} . \]  

(5-11c)

Thus, as one would expect, in the Anderson model the three parameters of \( T(N) \) are not independent, but all depend on the same random variable, \( \epsilon_N \), and their distributions (for fixed \( E \)) are determined by the rectangular distribution of the \( \epsilon_N \). The shape of the distributions of \( \phi \) and \( \phi \) are obvious from (5-11b) and (5-11c); the distribution of \( \theta \) is, from (5-11a)

\[ Q(\theta) d\theta = \begin{cases} 
\left( \frac{1-(E/2)^2}{W \cos^2 \theta} \right)^{1/2} d\theta & |\theta| \leq \theta_0 \\
0 & |\theta| > \theta_0 
\end{cases} \]  

(5-12)

where \( \theta_0 = \tan^{-1} \left[ \frac{W}{(4-E^2)^{1/2}} \right] \). Equations (5-3), (5-11), and (5-12) will be used to study \( P(\nu) \) in the weak and strong disorder limits.
In principle, the recursion relation (5-3b) determines \( v_N \) as a function of \( \{ \varepsilon_i \} \), and thus determines a definite distribution \( P_N(v_N) \) for a given distribution of the \( \varepsilon_i \). The hypothesis is that as \( N \to \infty \), \( P_N(v_n) \) converges to a fixed limiting distribution \( P(v) \) which is in fact uniform over the interval \([0,2\pi]\). It is, however, quite difficult to find \( P(v) \) from (5-3b) analytically in the weak disorder limit. However, it is easy to see that given some fixed distribution \( P_0(v_0) \) for a chain of length \( n_0 \), (5-3b) tends to randomize \( v_n \) over \([0,2\pi]\) for \( n > n_0 \).

Suppose \( n_0 >> 1 \), then \( \left( \frac{\rho_{n_0}}{4+\rho_{n_0}} \right)^{\frac{1}{2}} \approx 1 \), whereas, from (5-11b) we see that \( b_n = \left( \frac{z_n}{1+z_n} \right)^{\frac{1}{2}} \approx W \). So to lowest order in \( W \) (5-3b) gives us

\[
v_{n_0+1} \approx v_{n_0} + \delta_{n_0+1}.
\]

Iterating we find

\[
v_{n_0+n} = \sum_{i=1}^{n} \delta_{n_0+i} + v_{n_0} = \sum_{i=1}^{n} \theta_i + (\phi_{n_0+1} - \phi_{n_0+n}) + v_{n_0}. \tag{5-13}
\]

From (5-11c) we see that \( \phi_{n_0+1} - \phi_{n_0+n} \) equals \( 2kd \), or \( 2kd + \pi \), with equal probability; whereas the sum of the \( \theta_i \), which are each uniformly distributed from \(-W/(4-E^2)^{\frac{1}{2}}\) to \(W/(4-E^2)^{\frac{1}{2}}\) if \( W << 1 \), will become approximately normal for \( n >> 1 \). Thus, as \( n \) increases and the two gaussian peaks at \( 2kd \) and \( 2kd + \pi \) widen, the term added to \( v_{n_0} \) will become uniformly distributed and will tend to randomize \( v_{n_0+n} \) uniformly relative to \( v_{n_0} \). However, this linearized analysis only holds for \( n_0 >> 1 \) and \( nW << 1 \) and we shall see from the numerical results that it misses an essential feature of the approach of \( P(v) \) to uniformity, which is its dependence on energy.
Using Eqs. (3-6), (3-7), and (5-9) we numerically generated ensembles of chains of varying lengths and computed histograms of $P_N(v_N)$. It was immediately clear that for $E \neq 0$ and chains of reasonable length ($N \geq 1000$), $P_N(v_N)$ converged to a uniform distribution, $P(v) = 1/2\pi$. For $E = 0$, however, we found that the limiting distribution $P(v)$ was not uniform, but was moderately peaked at zero and $\pi$. Typical results for chains of 1000 atoms with $E \neq 0$ and 10,000 atoms with $E = 0$ are shown in Figs. 5-la and 5-lb. The fact that the limiting distribution $P(v)$ depended on energy was initially surprising since previous research\textsuperscript{20,34} had not indicated that the band-center state for diagonal disorder has any special properties. However, the deviations from uniformity even at $E = 0$ are small enough that $N\gamma_2$, the phase-dependent contribution to $\ln(1+P_N)$, is still negligible for long chains. Thus, the mean inverse localization length is determined solely by $\gamma_1$, as it is in the random phase model. Nonetheless, the special behavior at $E = 0$ provides a clue that there is a length scale in the model which depends on the electronic energy, and this is what our calculations find.

To determine the length scale at which phase randomness is reached we need to examine $P_N(v_N)$ as a function of $E$ and $W$. This requires an accurate statistical test for the uniformity of a sampling from a circular distribution. Such a test is provided by calculating the trigonometric moments of a sampling of $m$ values $v_i$ from $P(v)$, which are defined by

$$M_q = \frac{1}{m}\left(\sum_{i=1}^{m} \cos qv_i\right)^2 + \frac{1}{m}\left(\sum_{i=1}^{m} \sin qv_i\right)^2 , \quad (5-14)$$
where $q$ is a positive integer.\textsuperscript{41} It is not hard to show that for $m$ large, the quantity $T_p = \sum_{q=1}^{p} q M_q$ has a chi-squared distribution with $2p$ degrees of freedom if $P(v)$ is uniform.\textsuperscript{42} This test was found to provide more precise and consistent results than a Neyman-Pearson chi-squared test, and using it we were able to determine, for a given energy, the length at which $P(v)$ becomes uniform (with 95% confidence) to within 10-20 atoms.

We find that the length at which $P(v)$ becomes uniform is independent of the degree of disorder $W$ (as long as $W << 1$), and is inversely proportional to the energy. These results are illustrated in Figs. (5-2) and (5-3). Figs. (2a-2d) show that varying the disorder by two orders of magnitude has no effect on $P(v)$, which is highly non-uniform for chains of length 10 and uniform for chains of length 50. Hence the phase randomness length, $N_p$, is independent of the localization length. Fig. (5-3) shows that reducing the energy by a factor of five and increasing the chain length by an equivalent factor produces an essentially unchanged distribution $P(v)$, indicating that $N_p \propto E^{-1}$.

Why should $N_p$ depend on energy and not on $W$? Some insight into this may be obtained by examining Eqs. (5-11) as a function of energy. Nothing dramatic happens to $\theta_N$ at $E = 0$; it is approximately uniformly distributed with a width which varies slowly near $E = 0$. However, $\phi_N$ has a very special behavior at $E = 0$. At $E = 0$, $2Nkd = N\pi$ and $P(\phi_N)$ consists of two $\delta$-function peaks of equal weight centered at $\pi/2$ and $3\pi/2$ for each $N$. For $E \neq 0$, $P(\phi_N)$ still consists of 2 peaks separated by $\pi$, but the location of the peaks shifts with $N$. The shift at each
step is \((\pi - 2kd)\), so when \(N(\pi - 2kd) \approx 2\pi\) the peaks have traversed the interval \([0,2]\). If we hypothesize that this length is roughly \(N_p\) for energies near zero where \(E = \pi - 2kd\), then for \(E = 0.25\), this gives \(N_p \approx 25\), in good agreement with our numerical results which found \(P(\nu)\) failing a chi-squared test for uniformity at \(N = 20\) and passing at \(N = 30\) for \(E = 0.25\). Also, if the behavior of \(\phi_N\) is setting the length scale we can understand why \(N_p\) is independent of \(W\), since we see from (5-11b) that the distribution of \(\phi_N\) is entirely independent of \(W\). It depends only on the probabilities that \(\epsilon_n\) is positive or negative, which are, of course, equal for any degree of disorder. In summary, it seems quite plausible that the phase randomization length, \(N_p\), in the 1d Anderson model is determined by the extent to which the electron's wavelength is incommensurate with the lattice, i.e., it is the number of lattice spacings it takes for the phase of the electronic wave-function to make its closest approach to the value it had on the initial site.

Knowing that \(P(\nu)\) is uniform, we can now calculate the mean and variance of \(\ln(1+\rho_N)\) or, equivalently, the inverse localization length, which is equal to \((1/2N)\ln(1+\rho_N) = \gamma/2\) in 1d \(^{24}\) (this relationship will be discussed in detail in Chapter VII). Since \(P(\nu)\) is uniform, \(\langle\gamma_2\rangle = 0\) as noted above (we assume that the contribution to \(\gamma_2\) from the initial few atoms, where \(P(\nu)\) is not uniform, is negligible). Thus

\[
\langle\gamma\rangle = \langle\gamma_1\rangle = \frac{1}{2W} \int_{-W}^{W} \ln(1+\epsilon^2/4-E^2) d\epsilon = W^2/3(4-E^2) ,
\]

which agrees with the known result from perturbation theory. \(^{43}\) More-
over, we can also compute the variance of $\gamma$, which, as we will show in Chapter VI, should be proportional to $N^{-1}$.

$$\text{Var}(\gamma) = <\gamma^2> - <\gamma>^2 = <\gamma_1^2> + <\gamma_2^2> + 2<\gamma_1\gamma_2> - <\gamma_1>^2 \quad (5-16)$$

The first and last terms in (5-16) combine to give $\text{Var}(\gamma_1) = \frac{1}{N} \text{Var} [\ln(1+2)]$ which is of order $W^4/N$ and may be neglected. $<\gamma_1\gamma_2> = 0$ since $\gamma_1$ is phase-independent. So

$$\text{Var}(\gamma) = <\gamma_2^2> = \frac{1}{2NW} \int_{-W}^{W} d\epsilon \int_{0}^{2\pi} \frac{d\nu}{2\pi} \left[ \ln \left( 1 + \frac{\epsilon^2}{4-E^2} + 2 \left( \frac{\epsilon^2}{4-E^2} \right)^{\nu^2} \cos \nu \right) \right]^2$$

$$= \frac{2W^2}{3(4-E^2)N} + O(W^4/N) \quad (5-17)$$

Note, $(N/2)\text{var}(\gamma) = <\gamma>$; this relationship is significant since, if $\gamma$ is normally distributed, this implies that $\ln<\phi_N>/\ln<\phi_N> = 2$, which is a general result of the random phase model (see Eqs. (3-2), and (3-4)) and thus should hold here by virtue of the equivalence of the two models, which we have just established. This point, and other related issues will be discussed in detail in the following chapter.

In the limit of strong disorder ($W >> 1$) the recursion relation for $\nu_N$ obtained from (5-3b) is now tractable by the means employed by Lambert and Thorpe. In this limit, both $a_N$, $b_N \approx 1$ plus terms of order $W^{-1}$, so the right hand side of (5-3b) is unity, and we have

$$\nu_N = \theta_N + \theta_{N-1} - \phi_N + \phi_{N-1} \quad (5-18)$$
This is Eq. (10) of Ref. 19. Now $v_N$ only depends on the phases of the $N\text{th}$ and $(N-1)\text{th}$ barrier.

First, we analyze the contribution of the $\theta$ variables to $P(v_N)$. For large disorder the distribution $Q(\theta)$, given by Eq. (5-12) extends almost from $-\pi/2$ to $\pi/2$ and is strongly peaked at the ends, with a maximum value $W/(16-4E^2)^{1/2}$ a distance $(4-E^2)^{1/2}/W$ from either end, and a minimum value $(1-(E/2)^2)^{1/2}/W$ at the origin. The distribution of the variable $\mu_N = \theta_N + \theta_{N-1}$ is then determined by the convolution integral

$$D(\mu_N) = \frac{(1-(E/2)^2)^{1/2}}{W} \int_{-\infty}^{\infty} \frac{d\theta}{\cos^2 \theta \cos^2 (\mu_N - \theta)} \quad , \quad |\mu_N| < 2\theta_0$$

(5-19)

where $\theta_0 = \pi/2 - (4-E^2)^{1/2}/W$ for large $W$. This integral can be evaluated analytically in terms of elementary functions, but the result is rather cumbersome. Analysis of the exact result, however, gives the salient features of $D(\mu_N)$. It has a peak of order $W$ at zero, two symmetric peaks of order $W$ at $\pi \pm 2(4-E^2)^{1/2}/W$, and values of order $1/W$ at $\pm \pi/2$.

This is just what one would expect: all the values of $\theta$ are clustered near $\pm \pi/2$, so if $\epsilon_N$ and $\epsilon_{N-1}$ have the same sign, $\theta_N + \theta_{N-1} \approx \pi$, if they have opposite sign $\theta_N + \theta_{N-1} \approx 0$. The distribution falls off from its maxima roughly as $(\mu_N - \mu_{\text{max}})^{-2}$, because of the shape of $Q(\theta)$. Now $v_N \approx \delta_N = \mu_N - \phi_N + \phi_{N-1}$, and from Eq. (5-11b)

$$\phi_{N-1} - \phi_N = 2\cos^{-1}(E/2) + \pi/2 \left| \frac{\epsilon_N}{|\epsilon_N|} - \frac{\epsilon_{N-1}}{|\epsilon_{N-1}|} \right| \quad .$$

(5-20)

So, we see that if $\epsilon_N$ and $\epsilon_{N-1}$ have the same sign, then $\phi_{N-1} - \phi_N =$
 But if this is so then \( \mu_N \approx \pi \) with high probability, so \( \nu_N \approx \pi + 2\cos^{-1}(E/2) \) with high probability. If \( \varepsilon_N \) and \( \varepsilon_{N-1} \) have opposite sign, \( \phi_{N-1} - \phi_N = 2\cos^{-1}(E/2) + \pi, \mu_N \approx 0 \) and still \( \nu_N \approx \pi + 2\cos^{-1}(E/2) \). Thus, in either case we expect \( \nu_N \) sharply peaked around \( \mu_{\text{max}} = \pi + 2\cos^{-1}(E/2) \), and falling off as \( (\nu_N - \nu_{\text{max}})^{-2} \). This behavior is illustrated in Fig. (5-4). The shape of \( P(\nu) \) is independent of \( N \), so \( P(\nu) \) is never uniform over \( 2\pi \) for large disorder.

We can evaluate \( \langle \gamma_2 \rangle \) in the limit as \( W \to \infty \). At \( E = 0 \), \( \langle \gamma_2 \rangle \) is maximized as \( W \to \infty \), \( P(\nu) \) becomes a \( \delta \)-function at \( \nu = 0 \), and \( t_N \to 1 \), so

\[
\langle \gamma_2 \rangle = \int_{\nu \to \infty} P(\nu) \ln(2 + 2\cos \nu) = \ln 4 . \tag{5-21}
\]

Thus, \( \langle \gamma_2 \rangle \) is bounded by \( \ln 4 \) as \( W \to \infty \). On the other hand, evaluation of Eq. (5-15) in the limit \( W \to \infty \) shows that \( \langle \gamma_1 \rangle \propto \ln \left[ \frac{W^2}{4-E^2} \right] \), so

\[
\lim_{W \to \infty} \langle \gamma \rangle = \langle \gamma_1 \rangle = \ln \left[ \frac{W^2}{4-E^2} \right] . \tag{5-22}
\]

Again, Eq. (5-22) is in agreement with the known result, derived from perturbation theory.\(^{43}\) Since \( \langle \gamma \rangle \) only increases as \( \ln(W) \), for large finite \( W \), \( \langle \gamma_2 \rangle \) can make an appreciable contribution to \( \langle \gamma \rangle \), and our numerical calculations show that this is indeed the case.

To summarize, the phase distribution \( P(\nu) \) of the 1d Anderson model with diagonal disorder has the following properties: (1) In the limit of weak disorder \( P(\nu) \) is uniform over \([0,2\pi]\) for all \( E \neq 0 \) as \( N \to \infty \). (2) In this limit, \( P(\nu) \) is not uniform at each site, but becomes uniform at a length scale, \( N_p \) which is independent of disorder and is in general
much shorter than the localization length. (3) In the limit of large disorder $P(v)$ is independent of length and non-uniform. (4) In neither limit does the phase-dependent term, $\gamma_2$, make a major contribution to the resistance.

These results are almost exactly the opposite of those obtained by Lambert and Thorpe for the randomly spaced $\delta$-function model. In that model, they find that $P(v)$ is only uniform if $P(v)$ is uniform for one site; and that it is uniform at one site in the limit of strong disorder but not for weak disorder. It is easy to find the source of the very different behavior of the two models. In the randomly-spaced $\delta$-function model we have

$$
\frac{Z_n}{\mu_n} = \frac{\mu_n}{k}, \quad \theta_n = \tan^{-1}\left(\frac{\mu_n}{k}\right), \quad \phi_n = 2kx_n - \pi/2 ,
$$

where $u_n = u_0$ is the strength of the $\delta$-function, $m$ is the electron mass, $k$ is the wave-vector of the incident electron and $x_n$ is the position of the $n$th $\delta$-function. Since all the $\delta$-functions have the same strength, in this model, $Z_n$ and $\theta_n$ are not random at all. This has three immediate consequences. First, $\gamma_1 = \frac{1}{N} \sum_{n=1}^{N} \ln(1+Z_n)$ can never be zero no matter what the degree of disorder. Therefore, since $\ln(1+\rho_n)$ must be zero at zero disorder, $\gamma_2$ must converge to $(-\gamma_1)$ as the disorder goes to zero. Thus, in this limit $\gamma_2$ can never be negligible, whereas in the Anderson model $\gamma_1 \to 0$ as $W \to 0$, so $\gamma_1$ and $\gamma_2$ can go separately to zero with $\gamma_1$ always much greater than $\gamma_2$. Second, since $Z_n$ is non-random, it is possible in the $\delta$-function model to go to the strong scattering limit.
(Z_n \gg 1) without going to the limit of large disorder in the spacing of the δ-functions. For Z_n \gg 1, and n \gg 1, it is always true that v_n is approximately equal to the local phase \( \delta_n \) and therefore v_n can only be uniform if \( \delta_n \) is uniform. However, in the Anderson model \( Z_n \) is random and \( \langle Z_n \rangle \sim W^2 \), so one cannot go to the strong scattering limit without going to the strong disorder limit. There exists no limit analogous to the one in which the phase randomness assumption breaks down in the δ-function model. Finally, since the δ-function strength is always a positive constant, the phase \( \phi_n \) is not subject to a random phase shift of \( \pi \), which depends only on the sign of the scattering potential.

Any model with both positive and negative barriers will have this feature, and as discussed earlier, it appears to play a crucial role in phase randomization in the Anderson model. The physical origin of this random phase shift is most easily seen by considering scattering from a rectangular potential step of arbitrary sign. Here the continuity conditions require that the off-diagonal element of the transfer matrix be proportional to the quantity \( \eta = \left( \frac{k}{k'} - \frac{k'}{k} \right) \), where k is the electron wave-vector outside the step and k' is the wave-vector inside. If the step is positive, then \( \eta \) is positive and vice-versa, so the possibility of a \( \pi \) phase-shift in \( \phi \) is independent of the height of the step. Exactly the same thing will happen in the δ-function model if one allows the strength \( u_n \) to vary in sign. In fact, one can see from Eqs. (5-11) and (5-23) that the transfer matrix of the δ-function model with uniformly distributed random strengths \( u_n \) and non-random positions \( x_n \) is exactly equivalent to that of the Anderson model with diagonal disorder,
and the results of our work should hold for this model with the substitution $\epsilon_n/(4-E) + 2 \mu_n/k$. Thus, while the findings of Lambert and Thorpe concerning lack of phase randomization in the randomly space $\delta$-function model are correct, they are by no means typical of random 1d systems. They certainly do not apply to the 1d Anderson model. Moreover, our work indicates that any 1d model with a random potential which is positive or negative with equal probability will behave like the Anderson model, and be essentially equivalent to the random phase model in the limit of weak disorder. This will allow us to make some fairly general statements about the statistics of the resistance in 1d disordered systems in the next chapter.
Chapter VI. Strange Statistics

In this chapter, we finally confront directly questions concerning the statistical distribution of $\rho$ and related functions in disordered 1d systems. Let us summarize what we have established up to this point about the average behavior of $\rho$. From the recursive calculations of Chapter III we know that for the Anderson model with arbitrary disorder $\ln \langle 1+\rho \rangle \sim N$, and $(\text{Var}(1+\rho) / \langle 1+\rho \rangle^2) \sim e^N$ as $N \to \infty$. Our numerical calculations of Chapter IV indicate that $\langle \ln(1+\rho) \rangle \sim N$ for any energy and any combination of diagonal and off-diagonal disorder, except for purely off-diagonal disorder at $E = 0$. They also show that $\text{Var}[\ln(1+\rho)] / \langle \ln(1+\rho) \rangle^2 \sim N^{-1}$. In Chapter V, we showed that the Anderson model with weak diagonal disorder (and probably a much wider class of 1d models in the weak disorder limit) are equivalent to the random phase model, where $\langle \ln(1+\rho) \rangle$ is always proportional to the length of the chain. A hypothesis which explains all these results is that in all cases $\ln(1+\rho)$ is normally distributed for $N \gg 1$,

$$G(\ln(1+\rho)) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp \left[ - \frac{(\ln(1+\rho) - N\gamma)^2}{2N\sigma^2} \right] . \quad (6-1)$$

Thus $(1+\rho)$ has the lognormal distribution

$$P(1+\rho) = \frac{(1+\rho)^{-1}}{\sqrt{2\pi N\sigma^2}} \exp \left[ - \frac{(\ln(1+\rho) - N\gamma)^2}{2N\sigma^2} \right] . \quad (6-2)$$

By now we have good reason to believe (6-1) is approximately correct for any disordered 1d system, except at isolated singular points in
their spectrum; however, no entirely general proof of this exists.

We now present several arguments which prove (6-1) under various conditions.

First consider the limit of weak disorder, and assume we are treating a model (such as the Anderson model) where the phase randomness assumption holds. Note, by weak disorder we only mean that $R/T \ll 1$ for a single scatterer, not that $\ln(1+\rho)$ for the chain is small. Using the formalism of Chapter V (Eqns. (5-5) and (5-6)), we can write

\[
\ln(1+\rho_N) = \sum_{i=1}^{N} \ln(1+Z_i) + \sum_{i=2}^{N} \ln(1+t_i + 2t_i \cos \mu_i) \\
\approx 2 \sum_{i=1}^{N} (Z_i^2 \cos \mu_i + Z_i \sin^2 \mu_i)
\]

(6-3)

where we have expanded the logarithms to second order in the small quantity $Z_i^2$ (which is $\sqrt{R/T}$ for a single scatterer) and we have assumed $\rho_N/(1+\rho_N) \approx 1$ with very high probability.

In order to calculate the probability density of $\ln(1+\rho_N)$ for $N >> 1$ we begin by calculating its fourier transform (this is called the characteristic function of $\ln(1+\rho_N)$ in probability theory). This may written as

\[
C(k) = \left[ \prod_{i=1}^{N} \frac{dz_i}{z_i} P(z_i) \right] \left[ \prod_{i=2}^{N} \frac{d\mu_i}{2\pi} \right] \int dy \exp(iky) \\
\times \delta \left[ y - 2 \sum_{i} (Z_i^2 \cos \mu_i + Z_i \sin^2 \mu_i) \right]
\]

(6-4)
\[ C(k) = \left\{ \int dZ P(Z) \int \frac{d\mu}{2\pi} \exp\left[2ik(Z^2 \cos\mu + Z\sin^2\mu)\right] \right\}^N , \quad (6-5) \]

where \( P(Z) \) is some arbitrary probability density, and we have introduced explicitly the assumption that each \( \mu_i \) is uniformly distributed over \( 2\pi \) and independent of \( Z \). Since we are considering the weak disorder limit, \( \text{Prob}[Z << 1] \approx 1 \), and we may expand the exponent to get

\[ C(k) \approx \left[ 1 + ik\gamma - \frac{k^2\sigma^2}{2} \right]^N \quad (6-6a) \]

with

\[ \gamma = <Z> \quad (6-6b) \]
\[ \sigma^2 = 2<Z> = 2\gamma \quad , \quad (6-6c) \]

where we have integrated over \( \mu \) and kept only lowest order terms in \( <Z> \). Thus

\[ C(k) \approx \exp\left[ikN\gamma - \frac{k^2N\sigma^2}{2} \right] \quad . \quad (6-7) \]

Alert readers will recognize (6-7) as the fourier transform of a gaussian centered at \( N\gamma \); and indeed inverse transforming (6-7) yields exactly Eq. (6-1) for the distribution of \( \ln(1+p) \), with \( \gamma \) and \( \sigma^2 \) given by (6-6b,c). This proves that in the weak disorder limit (\( \gamma << 1, e^{N\gamma} > 1 \)) any model where the random phase assumption holds (e.g. the Anderson model) gives rise to a lognormal distribution of \( \rho \).

Now consider the opposite limit of large disorder (\( <Z> >> 1 \)); it is easy to show that \( \ln(1+p) = \ln\rho \) is normally distributed in the Anderson model. Abrahams and Stephen have shown this for very large diagonal
disorder and their proof can be trivially generalized to both types of disorder at $E = 0$. In this limit, the dominant contribution to $\rho$ is obtained by dropping the second term in the recursion relations for $(P^{11}_N)^2$ in (3-14).

$$
(P^{11}_N)^2 = \frac{(\epsilon_N)^2}{(V_{N,N+1})^2} (P^{11}_{N-1})^2,
$$

which implies

$$
\ln((P^{11}_N)^2) = \sum_{n=1}^{N} \ln \left( \frac{(\epsilon_n)^2}{V_{n,n+1}} \right),
$$

and, since $\ln \rho = \ln((P^{11}_N)^2$ in this approximation (see Eq. (3-13)), by the central limit theorem $\ln \rho$ is normally distributed and $P(\rho)$ is given by (6-2) with $\gamma = \langle \ln \epsilon^2 \rangle - \langle \ln V^2 \rangle$ and $\sigma^2 = \text{Var}[\ln(\epsilon^2/\nu^2)]$. Note, in this limit there is no special relation of $\sigma^2$ to $\gamma$. We have now proved that $\ln(1+\rho)$ is normally distributed according to (6-1) for the Anderson model with diagonal disorder in the limits of weak and strong disorder. That $\ln(1+\rho)$ is normally distributed has also been shown analytically by Mel'nikov, and by Kree and Schmid, for the gaussian white noise model in 1d; and, numerically, by Andereck and Abrahams, for a random $\delta$-function model.

More generally, for any disordered 1d system, it appears clear from the argument in Chapters III and V that one can always write $\rho_N$ as a quadratic function of the matrix elements of a product of a large member of random matrices with unit determinant. There exists a great
deal of sophisticated mathematical work on the probabilistic behavior of such products; and O'Connor has applied some of these ideas to prove a central limit theorem for a disordered classical harmonic chain. It seems quite possible that his approach can be generalized to prove that the logarithm of $\rho_N$ is normally distributed with mean growing as $N\gamma$ ($\gamma > 0$) for any model (as long as the transfer matrices are of a sufficiently general form). However, no such extremely general mathematical proof of this statement has actually been made; the best we can say is that a "virtual" proof of this theorem exists.

Informed by this "virtual" theorem, we can understand why, for the Anderson model, purely off-diagonal disorder at $E = 0$ is a special case. In this case, and only in this case, are the promotion matrices purely diagonal or purely off-diagonal for every chain in the ensemble (see Eq. (3-6)); so the random matrix products which can occur are restricted to an extremely small subspace of the probability space. It is not surprising that anomalous behavior can occur when such a severe further constraint is imposed on the probability measure of these matrix products. Even in this special case, Eq. (4-10) shows that the distribution of $\rho$ is still lognormal in the tail (for $\rho >> 1$), but with $\gamma = 0$. It is not correct however to say that this is precisely a special case of (6-2) with $\gamma = 0$, since $P(\rho)$ is not lognormal for $\rho \leq 1$.

A further important remark is that while for a given $N$, $\ln(1+\rho_N)$ may be normally distributed to a good approximation, it does not follow that $\rho_N$ will be lognormally distributed to a good approximation. Exponentiating the approximately normal variable $\ln(1+\rho_N)$ will amplify any
small deviations from a true gaussian behavior, so that \( P(\rho_N) \) may be poorly approximated by a lognormal distribution. This accounts for certain discrepancies that arise from trying to infer the values of the parameters of \( G(\ln(1+\rho)) \) from knowledge of the moments of \( \rho \).

Having recognized these subtleties, we can now ignore them for most practical purposes and assume that (6-1) and (6-2) hold quite generally. Assuming this we can easily calculate the \( r^{th} \) moment of \( \rho \) (for the remainder of the chapter we shall assume \( \rho = 1 + \rho \)). We find

\[
<\rho^r> = \exp[rN\gamma + \frac{1}{2} r^2 N\sigma^2]
\]

(6-10)

from (6-10) we see that

\[
\frac{\text{Var}(\rho^r)}{<\rho^r>^2} = e^{rN\sigma^2} - 1.
\]

(6-11)

So for large \( N \) each moment of this distribution is relatively broader than the previous, and, as we found in the recursive calculations, the fractional variance of \( \rho \) diverges exponentially as \( N \to \infty \). Since, trivially, \( <\ln\rho> = N\gamma \), comparing the arithmetic and geometric means gives

\[
<\rho> = e^{(\gamma + \frac{1}{2} \sigma^2)N},
\]

(6-12a)

\[
\bar{\rho} = e^{<\ln\rho>} = e^{\gamma N}.
\]

(6-12b)

So \( <\rho> \) is exponentially greater than \( \bar{\rho} \). In order for (6-12a) and (6-12b) to be consistent with the results of the random phase model, that \( <\rho> \) grows exactly twice as fast as \( \bar{\rho} \) (see Eqs. (3-2) and (3-4)), the rela-
tion \( \gamma = (1/2) \sigma^2 \) must hold, which fortunately we derived in Eqs. (6-6b) and (6-6c).

Clearly, \( \langle \rho \rangle \) is much too large to be typical of the distribution; the reason is that it is skewed by the very long, slowly decaying tail of \( P(\rho) \). We can show this explicitly by asking the question: What is the probability that \( \rho \) is greater than \( \langle \rho \rangle \)?

\[
\text{Prob}[\rho > \langle \rho \rangle] = \text{Prob}\left[ \ln \rho > N\gamma + \frac{1}{2} N\sigma^2 \right] \\
= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{N\sigma^2}/2} \exp[-y^2/2]dy = 1 - \text{erf}(\sqrt{N}\sigma/2).
\]

The error function goes to unity extremely rapidly for arguments greater than 1, \( [(1 - \text{erf}(5)) < 10^{-5}] \), so \( \text{Prob}[\rho > \langle \rho \rangle] \) is extremely small for \( N > 4/\sigma^2 \). Since \( \frac{1}{2} \sigma^2 = \gamma \) this condition is essentially that the chain be longer than the localization length; so we expect \( \langle \rho \rangle \) to be unrepresentative of the ensemble for all such chains. Since \( \langle \rho \rangle \) is skewed by very large, very improbable values of \( \rho \), a numerically generated ensemble average is likely to miss these values in the tail of the distribution entirely, and give very different results from an analytic calculation. We will return to this problem below.

First, let us ask a second question: Is \( \bar{\rho} \) truly representative of \( P(\rho) \)? The answer is yes in a certain sense, but no in another important sense.

\( \bar{\rho} = \exp[\ln \rho] \) is representative of \( P(\rho) \) in that it is both the median and modal (most probable) value of \( \rho \). It is also represen-
tative in the sense that, for any positive ε,

\[
\text{Prob}\left\{ e^{N(y+\varepsilon)} > \rho > e^{N(y-\varepsilon)} \right\} = 2 \operatorname{erf} \frac{\sqrt{N} \varepsilon}{\sigma^2} - 1 \approx 1 \quad , \quad (6-14)
\]

for \( N > \sigma^2/\varepsilon^2 \). However, these are not very strong statements statistically; in an important sense there may be no "typical" value of \( \rho \).

We can show this by asking how rapidly the value of \( \langle \rho \rangle \) and \( \bar{\rho} \), computed from a finite number of samples with resistance \( \rho \); will approach its true mathematical expectation. Thus, we define for a finite ensemble of \( m \) chains

\[
\bar{\rho} = \frac{1}{m} \sum_{i=1}^{m} \rho_i \quad , \quad (6-15a)
\]

\[
\bar{\rho} = \exp \left[ \frac{1}{m} \sum_{i=1}^{m} \ln \rho_i \right] \quad . \quad (6-15b)
\]

From (6-10) and (6-12) we can easily calculate the fractional variance of \( \bar{\rho} \) and \( \bar{\rho} \). We find

\[
\frac{\text{Var}(\bar{\rho})}{\langle \bar{\rho} \rangle^2} = \frac{1}{m} \frac{\text{Var}(\rho)}{\langle \rho \rangle^2} = \frac{e^{N\sigma^2}}{m} \quad , \quad (6-16a)
\]

\[
\frac{\text{Var}(\bar{\rho})}{\langle \bar{\rho} \rangle^2} = e^{(N/m)\sigma^2} \quad . \quad (6-16b)
\]

This is quite enlightening. It means that we will have to look at an ensemble of \( e^{N\sigma^2} \) systems before we can begin to be confident that \( \bar{\rho} \) is very near \( \langle \rho \rangle \); whereas we need sample only \( N\sigma^2 \) systems before \( \bar{\rho} \) begins
to approach $\overline{\rho}$ with high probability. So again we see that $\overline{\rho}$ is much better behaved than $<\rho>$. However, if we cannot make $m > N\sigma^2$, which may be difficult for long chains, then $\overline{\rho}$ will also be badly behaved in the sense that it is highly unlikely that two values of $\overline{\rho}$ taken from two particular samplings will be anywhere near one another. In this important sense there is no "typical" value of $\rho$ as $N \to \infty$.

At this point there is a natural tendency to say: This is all very well, but when a variable is exponentially large the obvious thing to do, both for a theorist and an experimentalist, is to study its logarithm. And if we look at $\ln \rho$, none of these statistical subtleties make a real difference, right? Wrong. As shown above, both $\langle \ln \rho \rangle$ and $\ln <\rho>$ are proportional to $N$, yet a computer simulation or an experiment, which measures $\ln \overline{\rho}$ (the finite ensemble average), will never see this behavior for chains longer than the localization length (even at $T = 0$). The reason for this has been mentioned above, and is illustrated by Fig. 4-4. Clearly, $\ln \overline{\rho}$ is not scaling linearly with $N$ is Fig. 4-4. For large $N$, a finite ensemble average misses the very large, very improbable values in the tail of $P(\rho)$ which contribute to $<\rho>$. Not only does this mean $\overline{\rho} \neq <\rho>$, it also means $\ln <\rho>$ is not a reasonable approximation to $\ln \overline{\rho}$, and neither is $\langle \ln \rho \rangle$, for finite $N$. Instead, since $P(\rho)$ broadens as $N$ increases, for small $N$ (in a sense to be defined below) $\ln \overline{\rho} \approx \ln <\rho>$, as $N \to \infty$, $\ln \overline{\rho} \to <\ln \rho>$, and in between $\ln \overline{\rho}$ has a well-defined non-linear dependence on $N$. We will derive this behavior analytically below, and show that it occurs whenever one is dealing statistically with the exponential of a well-behaved random variable. Thus, in particular,
the scaling law we derive will hold for any model of a disordered 1d system where the distribution of \( \ln\rho \), denoted now by \( H_N(\ln\rho) \) for generality, is well-behaved with mean and variance growing as \( N \), the length of the sample.

We begin by assuming that \( \ln\rho \) is distributed with mathematical expectation \( \langle \ln\rho \rangle = \gamma N \) and \( \text{Var}(\ln\rho) = \sigma^2 N \), thus the fractional variance of \( \ln\rho \) goes to zero as \( N \to \infty \). We have shown that these properties of \( H_N(\ln\rho) \) imply that \( \rho \) is poorly behaved in the senses discussed above as long as \( \sigma \sqrt{N} \gg 1 \). We note that it follows trivially that \( \ln g = 1/\rho \) has the distribution \( H_N(-\ln\rho) \), with \( \langle \ln g \rangle = -\gamma N \) and the same variance. Therefore, if \( H_N \) is symmetric around its mean, then each step in our subsequent argument also applies to \( \ln g \) with the simple replacement \( \gamma \to -\gamma \).

We are interested in the behavior of the random variable

\[
\tilde{\rho} = \frac{1}{m} \sum_{i=1}^{m} \rho_i,
\]

and in particular that of \( \ln\tilde{\rho} \) for \( \sigma \sqrt{N} \gg 1 \). It is very useful to define a new random variable, \( \delta \), so that

\[
\ln\rho_i = \gamma N + \sigma \sqrt{N} \delta_i.
\]  

(6-17)

Thus, \( \delta \) has a distribution of the same form as \( H_N(\ln\rho) \), but with mean zero and standard deviation unity, which we will denote as \( H(\delta) \).

Then \( \tilde{\rho} = \frac{1}{m} \sum_{i=1}^{m} \exp[\gamma N + \sigma \sqrt{N} \delta_i] \), and since exponentiation spreads out the values of the \( \rho_i \) very broadly, the sum of the \( \rho_i \) tends to be dominated by its maximum value, call it \( \rho_0 \equiv \exp[\gamma N + \sigma \sqrt{N} \xi] \). One can see this clearly by considering, for fixed \( m \)
Thus, we expect $\ln \rho$ to behave like $\ln(\rho_0/m)$. In particular, the expected value, $\langle \ln \rho \rangle$, satisfies the scaling law

$$
\langle \ln \rho \rangle = \langle \ln \rho_0 - \ln(m) \rangle = \gamma N + \sigma \langle \xi \rangle_m \sqrt{N} - \ln(m), \tag{6-19a}
$$

and, similarly,

$$
\langle \ln g \rangle = -\gamma N + \sigma \langle \xi \rangle_m \sqrt{N} - \ln(m). \tag{6-19b}
$$

Moreover, we do expect observations of $\ln \rho$ to be well represented by $\langle \ln \rho \rangle$, since $(\ln \rho)^2 = (\ln \rho_0)^2$ implies $\text{Var}(\ln \rho) \propto N$ so that the fractional variance of $\ln \rho$ decreases as $1/N$.

Let us first examine the qualitative consequences of Eq. (6-19). First, it rigorously shows, as suggested by Anderson et al., \(^{15}\) that at zero temperature as $N \to \infty$ the scale resistance, $\exp[\langle \ln \rho \rangle] = e^{\gamma N}$, and conductance, $e^{-\gamma N}$, are the relevant quantities. From (6-19) we see that a finite ensemble average of $\rho$ or $g$ yields

$$
\frac{\ln \rho}{\langle \ln \rho \rangle} = \frac{\ln \rho}{\gamma N} = -\frac{\ln g}{\gamma N} \to 1 \text{ as } N \to \infty.
$$

Second, for large finite $N$, it predicts a correction to the scale resistance proportional to $\sqrt{N}$, which we shall see is important for explaining numerical results and may be experimentally observable under conditions to be discussed below. This $\sqrt{N}$ correction implies slower than linear growth of $\ln \rho$ (i.e. second derivative negative) for finite $N$, which has been observed in the numerical studies cited above; whereas all known analytic calcu-
lations give \( \ln \langle p \rangle \propto N \) for large finite \( N \). In particular, in the special case of the Anderson model with purely off-diagonal disorder at \( E = 0 \), it has been shown above that the parameter \( \gamma = 0 \). Thus (6-14) implies \( \ln \delta \propto \sqrt{N} \) in this case as was found in the numerical studies of Economou and Soukoulis,\(^{35}\) despite the fact that even in this special case the analytic result is \( \ln \langle p \rangle \propto N \).

Equation (6-19) has no adjustable parameters since the quantities \( \gamma, \sigma \) and \( \langle \xi \rangle \_m \) are all determined by the distribution of \( \ln p \) and the ensemble size. The parameter \( \gamma \) may be shown to be equal to twice the average inverse localization length, and for the Anderson model with diagonal disorder this is known from a perturbative calculation\(^{48}\) which gives \( \gamma = 0.084 (W/V)^2 \) at \( E = 0 \), where \( 2W \) is the rectangular distribution of site energies. In addition, we have just shown that many models of disordered 1d systems satisfy the relation \( \gamma = \frac{1}{2} \sigma^2 \) for weak disorder, including the Anderson model. Thus, in many cases Eq. (6-19) may be calculated entirely analytically, given the value of \( \langle \xi \rangle \_m \). Therefore, we should be able to get meaningful quantitative comparisons of (6-19) with numerical studies by taking the analytically or numerically determined values of \( \sigma \) and \( \gamma \), and by calculating \( \langle \xi \rangle \_m \). \( \langle \xi \rangle \_m \) is the expected value of the maximum of a set of \( m \) independent random variables with the standard distribution \( H(\delta) \). We shall find that in fact \( \langle \xi \rangle \_m \) is both relatively insensitive to the exact nature of \( H(\delta) \) and is a slowly varying function of \( m \).

To obtain the probability density for \( \xi \), \( P(\xi) \), we note that since \( \xi = \max(\delta_1, \ldots, \delta_m) \) it follows that \( \text{Prob}\{ \xi \leq Z \} = \prod_{i=1}^{m} \text{Prob}\{ \delta_i \leq Z \} \).
Therefore,
\[
P(\xi) \, d\xi = m[H(\xi)]^{m-1} \, dH.
\]  
(6-20)

It turns out to be easier to get a good approximation for \( \langle \xi \rangle \_m \) in general by considering it as a function of the random variable \( H \), so that
\[
\langle \xi \rangle \_m = m \int \! \xi(H) \, H^{m-1} \, dH.
\]  
(6-21)

The reason that this is convenient is that \( \langle H \rangle = \frac{m}{m+1} \) and
\[
\text{Var}(H) = \frac{m}{(m+1)^2 (m+2)}.
\]
so that for \( m \) large, \( \text{Var}(H) / \langle H \rangle^2 = 1/m^2 \) and \( \langle \xi(H) \rangle \) is close to \( \xi(\langle H \rangle) \). We can then Taylor expand \( \xi(H) \) around \( \langle H \rangle \) in (6-21) to give
\[
\langle \xi \rangle = m \int \! H^{m-1} \, dH \left[ \xi(\langle H \rangle) + \frac{\partial \xi}{\partial H} \bigg|_{\langle H \rangle} (H-\langle H \rangle) + \frac{1}{2} \frac{\partial^2 \xi}{\partial H^2} \bigg|_{\langle H \rangle} (H-\langle H \rangle)^2 + \ldots \right] \]  
(6-22)

\[
\langle \xi \rangle \_m = \xi(\langle H \rangle) + \frac{1}{2m} \xi''(\langle H \rangle) \]  
(6-23)

Equation (6-22) is an entirely general approximation for \( \langle \xi \rangle \) which is expected to be good for \( m >> 1 \), and whose accuracy can be made arbitrarily good by including higher terms in the Taylor series. All one needs to evaluate (6-22) is an expression for the inverse of the standardized distribution function \( H \) (note, \( H \) is not the probability density).

The case of particular interest to us is when \( \ln p \) is normally distributed (Eq. (6-3)), since we have argued that this is generally the case for disordered 1d systems. In this case \( H \) is the standard error
function which is approximated by $H(\xi) = 1 - \frac{\exp[-\xi^2/2]}{\sqrt{2\pi}\xi^2}$, for $\xi \gg 1$.

Then

$$\xi = \left\{ \frac{1}{2\pi (1-H)^2} - \ln(\xi^2) \right\}^{1/2}$$

and

$$\xi = \left\{ \frac{1}{2\pi (1-H)^2} - \ln\left[ \frac{1}{2\pi (1-H)^2} \right] \right\}^{1/2}$$

(6-24)

evaluating (6-24) at $<H> = m/(m+1)$ gives the first term in (6-23) and higher terms may be obtained straightforwardly. For $m = 1000$, $\xi(<H>) = 3.08$ and the next term, $\frac{1}{2m^2} \xi''(<H>) = 0.16$, so our series approximation for $<\xi>$ seems nicely converged. For a bounded distribution $H$ the calculation of $<\xi>$ is trivial for $m$ large since the probability that the maximum of $\{\delta\}$ is very near the upper bound becomes very high, and one can take $<\xi>$ equal to the upper bound. For a rectangular distribution with unit variance the upper bound is $\sqrt{3}$ and the exact result is $<\xi> = \sqrt{3} \frac{(m-1)}{(m+1)}$. Thus, as noted above, we see that for reasonable size ensembles ($m \approx 1,000 - 10,000$), $<\xi>$ only varies by a factor of two or three between these two very different distributions and in both cases $<\xi>$ is a slowly varying function of $m$.

Having calculated $<\xi>_m$ we are in a position to apply Eq. (6-19) to various numerical results. Now we assume that $\ln p$ is normally distributed and take $<\xi>_m$ from Eqs. (6-23) and (6-24). Figs. (6-1a) and (6-1b) show that Eq. (6-19) does indeed give good quantitative agreement with numerical data. In Fig. (6-1) the circles are the results of the computer experiment, while the thick curve is the analytic result for
ln<\rho> derived in Chapter III, which, as usual, predicts much too large values of ln\bar{\rho}. The agreement is excellent in view of the facts that:

(i) Eq. (6-19) is applied with no parameters left free; (ii) our computer experiment only constitutes one observation of ln\bar{\rho} and thus should have some small random deviation from the exact <ln\bar{\rho}>. In plotting Eq. (6-19), we have used the average value of the numerically obtained values of \sigma and \gamma for each length, which are \gamma = 0.0381, \sigma = 0.272. Using the perturbative calculation for \gamma, and the relation \gamma = 1/2\sigma^2 yields \gamma = 0.041, \sigma = 0.29, which gives a reasonable, but less good, fit to the data. This is to be expected, since the approximation of weak disorder is beginning to break down for W/V = 1/\sqrt{2}. Equation (6-19) also gives a good fit for the data in Fig. 4-4 using the numerically obtained values \gamma = 0.0163, \sigma = 0.196.

These results show that Eq. (6-19), and the assumption on which it is based, provide a useful description of computer experiments. Whether it will contribute to a useful description of actual experiments which measure the resistance of very thin metallic wires is a much more complicated question. In order to answer it we must first discuss the physics of conduction in disordered systems at finite temperature. And then determine what effect the enormous intrinsic statistical fluctuations of the zero-temperature transport coefficients may have on finite-temperature conduction. This is the subject of the next chapter.
Chapter VII. Conduction at Finite Temperature: The Role of Statistical Fluctuations.

In this chapter we begin a treatment of localization effects as they manifest themselves at finite temperature in highly disordered thin metallic wires. All the theoretical calculations presented up to this point have been done for a disordered system at zero temperature, as, indeed, are all the first principles calculations in the existing literature on localization in any dimension. There exists no rigorous, physically meaningful theory of electronic conduction in the presence of a random potential at finite temperature. The reason for this is fairly subtle; it is because at finite temperature the dynamic interactions between the electron and the lattice are essential to the physics of the problem. As noted in Chapter II, one can easily calculate the thermally averaged DC conductance for a system of non-interacting electrons in a random, static external potential, but this is a physically meaningless quantity. The dynamic interactions with the lattice introduce a time scale (and therefore a length scale, as we shall see) beyond which calculations based on a static, random potential are invalid. Thus, the finite temperature behavior of such disordered conductors is characterized by a complete breakdown of one of the fundamental tools of solid state theory, the Born-Oppenheimer approximation. This makes the theory of their finite temperature behavior both very interesting and very difficult.

Despite the lack of any rigorous finite temperature theory, there does exist a reasonable qualitative understanding of the physics of
finite temperature conduction in highly disordered metals, which is due primarily to Thouless and Mott. The basic picture is the following: we imagine the electron in a disordered metal placed on a particular site \( \mathbf{r} = 0 \) in the solid at time \( t = 0 \). We assume all the electronic states are localized, so this initial state is a superposition of localized eigenstates peaked near \( \mathbf{r} = 0 \). Since the initial state is not an eigenstate, the system begins a non-trivial time evolution governed simply by the time-dependent Schrödinger equation with a random potential. We expect that initially this wave-packet of localized states will spread out with \( \langle r^2 \rangle = D t \), which corresponds classically to a random walk of step length \( \ell = v_f \tau \) (\( D = \ell^2 / d \tau \) is the elastic diffusion constant in \( d \) dimensions, \( \ell \) is the elastic mean free path, \( \tau \) is the elastic scattering time, and \( v_f \) is the fermi velocity). However, the localization results at \( T = 0 \) are telling us that quantum mechanically, a monoenergetic random walk cannot last forever. When \( D t \approx \lambda^2 \), where \( \lambda \) is the localization length, then \( \langle r^2 \rangle \) will begin to increase much more slowly with time, in fact exponentially slowly, \( \langle r^2 \rangle \propto \ln t \). A quantum-mechanical drunkard can only stagger so far without a pick-me-up. However, at finite temperature this picture breaks down, because we cannot treat the electron as being in a fixed initial state which simply evolves according to the time-dependent single-particle Schrödinger equation. Instead, interactions with the thermal reservoir will cause the electronic state to change by an average energy \( k_B T \) in a time \( \tau_{\text{in}} \), the inelastic scattering time. On this time scale, the electron has lost all knowledge that it was initially placed at \( \mathbf{r} = 0 \),
and now may have its peak density essentially anywhere within a radius \( \lambda \) of its initial site. Thus, thermal effects cause a random walk on a different time scale \( \tau_{in} \) (instead of \( \tau \)) which is not affected by the destructive interference which traps the electron within a localization length of its initial site at \( T = 0 \). The resulting motion of the electron in linear response to an external field depends on the relative magnitude of the time and distance scales we have introduced above.

Consider cooling the system down from high temperature. At temperatures where \( \tau_{in} \) is on the order of \( \tau \) the electron never has enough time to travel a distance of order \( \lambda \) by repeated elastic scattering and hence never "knows" that it would be localized at \( T = 0 \). At these temperatures the conductance is just given by Ohm's law,

\[
g = \sigma_0(T) L d^{-2}, \quad \text{with} \quad \sigma_0(T) = \frac{ne^2 \tau_M}{m}, \quad \text{and} \quad \tau_M(T), \quad \text{the total scattering time, given by Mathiessen's rule,} \quad [\tau_M(T)]^{-1} = 1/\tau + 1/\tau_{in}.
\]

In this Ohmic regime the resistance is a monotonically increasing function of \( T \), since \( \tau_{in} \) is a monotonically decreasing function of \( T \). This is the regime where semi-classical transport theory holds, as embodied in the Boltzmann equation, and where thermal effects impede transport.

As we lower the temperature the electron begins to diffuse elastically farther and farther before scattering inelastically. The length it typically travels is the inelastic diffusion length,

\[
\xi_i = \sqrt{D\tau_{in}}.
\]

We enter what is called the weakly localized regime when \( \xi_i \gg \lambda \), and \( \xi_i \) approaches, but is less than \( \lambda \), the localization length. In this regime, one expects the resistance to reach a minimum value near the zero-temperature Ohmic resistance, \( r_0 \), and then begin to increase as
the temperature decreases. Following the suggestion of Thouless\textsuperscript{5,50}, one can obtain positive, temperature-dependent corrections to the Ohmic resistance in one and two dimensions by just replacing \( L \), the sample length, in the zero temperature scaling laws by \( \ell_1(T) \), which is an "effective" sample length. For example, in 1\( d \), assuming the random phase model is correct, we can obtain the resistance of a sample of length \( L \) at finite temperature by assuming that it is the sum of the resistances of \( L/\ell_1 \) segments of length \( \ell_1(T) < \lambda \). Using Eq. (3-3) we can estimate that

\[
r = \frac{\langle n \rangle}{e^2} \frac{(L/\ell_1)}{\ell_1} \left[ \exp \left( \frac{2 \ell_1}{\lambda} \right) - 1 \right] = \frac{2 \pi \hbar L}{e^2 \lambda} \left( 1 + \frac{\ell_1}{\lambda} \right). \quad (7-1)
\]

For this recipe to make sense, the first term in (7-1) must be the Ohmic resistance. If we regard our system as truly one-dimensional, then we can rewrite the first term as

\[
r_0 = \frac{m}{ne^2 \tau} \left( \frac{2 \lambda}{\lambda} \right) L, \quad (7-2)
\]

where we have used the fact that in 1\( d \), in the weakly disordered regime, the single-spin number density is \( n = k_f / \pi \). Of course, the factor \((m/ne^2 \tau) = (\sigma_0)^{-1}\) is just the \( T = 0 \) Ohmic resistivity, so for equation (7-2) to give Ohm's law for the resistance we require that \( \lambda \), the localization length in 1\( d \), be equal to twice the elastic mean free path. Fortunately, Thouless has shown in a quite different way that precisely this identity is satisfied.\textsuperscript{52}
We wish to apply (7-1) to a 3-dimensional thin wire. Since for such a system, there is some ambiguity in defining the localization length \( \lambda \) which appears in (7-1), we will simply adopt the definition

\[
\lambda = \frac{2\pi \hbar}{e^2} \sigma_0 A .
\]  

This merely says that \( \lambda \) is the length of the wire such that it has an Ohmic resistance equal to \( 2\pi \hbar/e^2 \) (\( \sigma_0 \) is now the Ohmic conductivity in 3d, and \( A \) is the cross-sectional area of the wire). With this interpretation (7-1) predicts a temperature-dependent increment to the Ohmic resistance, \( r_0 \), in the weakly localized regime of the form

\[
\frac{\Delta r}{r_0} = \frac{\varepsilon}{\lambda} = \left( \frac{e^2}{2\pi \hbar} \right) \frac{\sqrt{D_{\text{in}} \ln}}{\sigma_0 A} .
\]  

In principle, all the parameters in (7-4) are known or can be estimated for a particular material, however, in practice determining \( \tau_{\text{in}}(T) \) either experimentally or theoretically is quite difficult as we shall see in the next chapter. Nonetheless, we do expect that typically \( \tau_{\text{in}}(T) \propto T^{-p} \), where \( p \) is of order unity; so (7-4) predicts a slowly increasing increment to \( r_0 \) in the weakly localized regime as \( T \rightarrow 0 \).

There have been a large number of recent experiments which attempt to measure such an effect in thin metallic wires,\(^{8,53-58}\) and there is no doubt that such behavior occurs in real systems. However, the interpretation of these experiments has been complicated by certain discrepancies between theory and experiment, as well as by rival theoretical predictions. These issues will be discussed in detail in the following chapter.
The behavior predicted by Eqs. (7-1) and (7-4) does not, of course, persist down to zero temperature; the motion of the electron changes qualitatively once \( \xi \) becomes greater than \( \lambda \). In this intermediately localized regime the electron typically diffuses a distance \( \lambda \) in a time \( t_\lambda = \lambda^2 / D \), and since \( t_\lambda \ll \tau_{in} \), it must wait for a thermal interaction to move much further. The total diffusion constant for this process (not to be confused with the elastic diffusion constant \( D \), which describes Ohmic transport in the absence of localization), has been estimated by Thouless\(^50\) to be \( D_{tot} = \lambda^2 / 12\tau_{in} \). Using the Einstein relation, we find in this regime a conductivity,

\[
\sigma = \frac{e^2 k_f^2 \lambda^2}{96\pi K v_f \tau_{in}} \quad ,
\]

and therefore a resistance, for a segment of length \( \lambda \)

\[
r(\lambda) = \frac{96\pi K v_f \tau_{in}}{e^2 k_f^2 \lambda A} .
\]

Note, now the resistance varies as \( \tau_{in} \), not as \( \sqrt{\tau_{in}} \), as it does in the weakly localized regime. Thus, one should see the resistance begin to increase more rapidly with decreasing temperature at low enough temperatures that \( \xi > \lambda \). It appears possible in principle to see this qualitative change experimentally, but so far experiments have apparently not probed low enough temperatures to do so. A more rigorous theory of this temperature regime would also be quite useful.
The next qualitative change in the temperature-dependence of the resistance should occur when the temperature is so low that the energy level spacing, \( \Delta E \), in a volume \( \lambda^3 \) is greater than \( k_B T \). Then the electron is in a bind. If it waits within \( \lambda \) of its initial site the probability of a thermal fluctuation of sufficient energy to delocalize it is proportional to \( e^{-\Delta E / k_B T} \), which is small. On the other hand, it can only sample larger regions of space at an exponentially slow rate. This competition leads to conduction by the familiar Mott variable-range hopping mechanism, which predicts

\[
 r_h \propto \exp \left( \frac{T_0}{T} \right)^{1/d+1} .
\]  

(7-7)

For a one-dimensional system, this would predict \( r \propto \exp[(T_0/T)^{1/2}] \); however, Kurkijarvi has argued that in 1d hops longer than \( \lambda \) cannot contribute to conduction and that therefore, in the strongly localized regime, the resistance of a wire will vary as

\[
 r_{1d}(T) \propto \exp[\frac{T_0}{T}] .
\]  

(7-8)

Fowler has found experimental evidence for (7-7) in 1d, and Kurkijarvi's argument for (7-8) is by no means clearly correct; however, for our present purposes, this makes little difference, and we will tentatively accept (7-8).

Our present purpose, as you may well have forgotten by now, is to assess the effect of intrinsic statistical fluctuations on the finite temperature transport coefficients. Having now elucidated the various
temperature regimes of interest, we can answer this question rather easily.

The important observation is the following: Although the calculations based on the Landauer formula of Chapters III–VI do not directly tell us the statistical behavior of the resistance at finite temperature, they do tell us the statistical behavior of the localization length. The inverse localization length, as mentioned earlier is just

\[ \gamma = (1/2L) \ln(1+p) \]

and this is a temperature-independent quantity. Since the finite temperature resistance, in each of the three regimes discussed above, only depends on the microscopic, quenched disorder through its dependence on the localization length, we can use the knowledge that \( \ln(1+p) \) is normally distributed at zero temperature to estimate the intrinsic statistical fluctuations in the resistance at finite temperature.

Up to this point we have been using the symbol \( \gamma \) to denote both the inverse localization length of a particular chain and the average inverse localization length of an ensemble of chains. Now it is convenient to denote the former quantity, which is a random variable, by \( \alpha \) and reserve \( \gamma \) for the latter, so \( \gamma \equiv \langle \alpha \rangle \) from now on, and

\[ \alpha = (1/2L) \ln(1+p) \]  

(7-9)

Then the arguments which establish the general validity of Eq. (6-1) also imply that \( \alpha \) is normally distributed with

\[ P(\alpha) = \frac{1}{\sqrt{4\pi \gamma/L}} \exp\left[ -\frac{(\alpha-\gamma)^2}{2(2\gamma/L)} \right] , \]  

(7-10)
where we have assumed \( \gamma = (1/2)\sigma^2 \) for convenience.

Using (7-10) we can now estimate the magnitude of the statistical fluctuations in the finite temperature resistance in the three regimes described above. In the weakly localized regime, \( \xi_i < \lambda \) and electron localization only causes a small correction to the Ohmic resistance. We now give a slightly more careful analysis of the resistance in this regime than we gave in (7-1), although given the non-rigorous nature of the arguments leading to (7-1), the following calculations should only be considered as order of magnitude estimates. Assume the sample consists of \( L/\xi_i \) segments of length \( \xi_i \) (we will ignore fluctuations in \( \xi_i^{61} \)), then the random variable, \( r \), which is the total resistance of the sample is given by

\[
r = \frac{L/\xi_i}{e^2} \sum_{n=1}^{L/\xi_i} \left( \exp(2\alpha_n \xi_i) - 1 \right),
\]

(7-11)

where \( \alpha_n \) is the inverse localization length of the \( n^{th} \) segment. So

\[
r \approx \frac{2\pi \hbar^2}{e^2} \sum_{n} \left( \alpha_n \xi_i + (\alpha_n \xi_i)^2 \right),
\]

(7-12)

and, in units of \( \frac{2\pi \hbar}{e^2} \)

\[
\Delta r = \sum_{n} (\alpha_n \xi_i)^2.
\]

(7-13)

If each \( \alpha_n \) is independently normally distributed according to (7-10), then (7-13) implies
Typical experimental results at temperatures of $1-10^6$K and for samples of length 500 microns find $\Delta r/r_0 = 0.01$. Taking $\gamma L_i = 0.01$, and $\ell_1$ on the order of a micron, Eq. (7-14) predicts variations in $\Delta r$ of only about 5% between two wires of the same length due to intrinsic fluctuations in the microscopic configurations of the wires. We, therefore, do not expect these fluctuations to be significant relative to other sources of variation in experiments on thin wires in the weakly localized regime; and this is the only regime these experiments have yet been able to probe.

In the intermediately localized regime ($\ell_1 > \lambda$) the situation is not much different. One can estimate $\lambda$ for a wire using the condition that $\lambda$ is the length of wire of cross-section $A$ which has an Ohmic resistance $2\pi\hbar/e^2$ (Eq. (7-3)) and one finds that for wires with $\sqrt{A} = 250\,\mu$A, $\lambda$ is not shorter than $\ell_1$ until temperatures of about $10^{-2}$K are reached (this assumes the slow variation of the inelastic scattering rate, $\tau_{in}(T) \propto 1/T$, which is found experimentally). In this regime, Eq. (7-6) tells us that the resistance is simply proportional to $\alpha$ and that

$$\frac{\text{rms}(\Delta r)}{\langle \Delta r \rangle} = \left[ 12(\gamma L_i) \frac{\ell_1}{L} \right]^{1/2}, \quad (7-14)$$

which again predicts an intrinsic variation of only a few percent.

It is only in the strongly localized regime, where the resistance
depends exponentially on the inverse localization length, that intrinsic fluctuations can become extremely large. Accepting Kurkijävi's argument, the resistance is given by a thermally activated hopping law of the form (7-8) with $K_B T_0 = \left(4\alpha/\pi n_f\right)$, where $n_f$ is the density of states per unit volume at the fermi level. Equation (7-10) then implies that

$$\langle r \rangle \sim \exp\left[\frac{d^2 \gamma}{L}\right],$$

with $d(T) = \left(4n_f A K_B T\right)^{-1} = \left(\frac{\pi^2}{2A k_f^2}\right)\left(\frac{\varepsilon_f}{K_B T}\right)$. Thus, the condition for fluctuations to become significant is

$$\frac{2 A^3 k_f^2 \varepsilon_f L}{3\pi^6} \left(\frac{K_B T^2}{\varepsilon_f}\right) \leq 1.$$  

For typical samples, with $A = 2 \times 10^{-11}\text{cm}^2$, $\ell = 5 \times 10^{-8}\text{cm}$, $L = 2 \times 10^{-2}\text{cm}$, Eq. (7-17) predicts that exponentially large fluctuations will only be visible at hopelessly low temperatures of around $10^{-6}\text{K}$. Note, however, that as we approach a truly one-dimensional system, with $A k_f^2 \to 1$ and $1/\gamma \to \ell$, equation is satisfied at about 10-50°K. Finally, when $d(T)$ becomes greater than $L$, we do expect to see a temperature-independent DC resistance which depends exponentially on $L$ and has enormous fluctuations from sample to sample. Even for a truly one-dimensional sample of length 100 microns this does not occur until $10^{-1}\text{K}$. We conclude that present experimental techniques will
not permit observation of the exponentially large fluctuations in the resistance due to fluctuations in the localization length. However, if progress is made towards fabricating an effectively one-dimensional system these effects may become visible, but only in the exponential hopping regime.

In the exponential hopping regime, it is interesting to examine the different scaling behavior of the mean resistance and the finite ensemble average resistance discussed in the previous chapter. We can rewrite Eq. (7-8) in the form \( r = \exp[\alpha d(T)] \), and then we have exactly the situation discussed in Chapter VI, the hopping resistance depends exponentially on a normally distributed random variable. The analysis of this statistical problem, which was given in detail earlier, (see Eq. (6-19)), implies

\[
\ln\langle r \rangle = \gamma d[1 + 2d/L] ,
\]

\[
\ln(r) = \gamma d[1 + \langle \xi \rangle (\frac{2}{\gamma L})^{\frac{1}{2}}] - \ln(\gamma L) ,
\]

where we have assumed the hopping resistance of the sample is roughly the sum of the resistances due to \( L/\lambda \) segments of length \( \lambda \). The interesting point which emerges is that as long as \( \lambda \ll d(T) \ll L \), then\n\[
\ln\langle r \rangle \approx \ln(\bar{r}) = \gamma d; \text{ thus, we see that statistical fluctuations only give a relatively small correction to the finite temperature resistance even in the exponential hopping regime, as long as } d(T) \ll L. \text{ Note also that } \gamma d \text{ is equal to } \langle \ln r \rangle \text{ and is the finite temperature analogue of the logarithm of the scale resistance, so in this regime this is always the}
relevant quantity. Finally, if \( d(T) \) becomes greater than \( L \), then the conductivity due to tunneling becomes greater than that due to hopping and all the results of the zero-temperature scaling theories should apply. Then we would expect the resistance averaged over many samples to satisfy Eq. (6-19). However, again, even for a truly one-dimensional wire of length 100 microns \( (A k T^2 \rightarrow 1) \), \( d(T) \) becomes greater than \( L \) only at very low temperature, about \( 10^{-10} \)K.

To summarize, at zero temperature the resistance is subject to exponentially large statistical fluctuations because it depends exponentially on a well-behaved random variable, the inverse localization length. At finite temperature inelastic scattering provides what are essentially alternate conducting "channels" to the zero temperature (tunneling) "channel", and these "channels" have much higher conductivity, so they dominate conduction. Only in the strongly localized regime, when the energy level spacing in a volume \( A \lambda \) is greater than \( K_B T \), will these inelastic channels have a resistance which depends exponentially on the inverse localization length. Thus, it is only in this regime that large intrinsic fluctuations in the resistance are to be expected, and these will only occur for a range of system parameters which are as yet experimentally inaccessible.
VIII. Inelastic Scattering Models for Weak Localization Effects in Thin Wires.

"Nothing is so simple that it can't be done with Green's functions." - M. Kastner.

In this chapter we introduce some new formalism and (hopefully) some new physical ideas in order to make an explicit, quantitative model for weak localization effects in thin metallic wires at finite temperature. In Chapter VII we presented a physical picture of the weakly localized regime, where \( \xi_1 < \lambda \), so that the electron typically scatters inelastically before diffusing a localization length. We found that in this regime the resistance is still approximately Ohmic, but it begins to increase as \( T \) decreases with a correction term proportional to the ratio \( \xi_1/\lambda \). As shown in the previous chapter, with some simple assumptions we can write this temperature-dependent increment to the Ohmic resistance \( r_0 \) in terms of experimentally measurable quantities (Eq. (7-4)),

\[
\frac{\Delta r}{r_0} = \frac{e^2}{2\pi \kappa} \frac{\sqrt{D \tau_{in}}}{\sigma_0 A} .
\] (8-1)

If we now substitute \( D = \frac{v_F \tau}{3} \), \( \sigma_0 = \frac{n e^2 \tau}{m} \) and replace the resistance scale \( \frac{2\pi \kappa}{e^2} \) by a somewhat flexible parameter, \( R_T \), we arrive at the form used by Giordano.²
\[ \frac{\Delta r}{r_0} = \frac{V_f}{eA\tau} \left[ \frac{2m}{3n} \rho_e \tau_{\text{in}}(T) \right]^{1/2}, \]  

(8-2)

where \( \rho_e \equiv 1/\sigma_0 \) is the Ohmic resistivity of the bulk material (we apologize for the slight clash with previous notation where \( \rho \) was always a resistance). Experimentally, one can easily measure the dependence of \( \Delta r/r_0 \) on area and temperature, and, with more difficulty, the dependence on \( \rho_e \). However, the theoretically predicted dependences, based on (8-2), are ambiguous, since \( \tau_{\text{in}}(T) \) in a dirty material can depend in an unknown way on \( A, T \) and \( \rho_e \). Thus, to get an explicit prediction out of (8-2) one must try to find the dominant inelastic scattering mechanism in such materials, and calculate \( \tau_{\text{in}}(T) \). Most experiments\(^8\) have found the results \( \Delta r/r_0 \) proportional to \( A^{-1} \) and \( T^{-1/2} \), and Giordano, in several experiments\(^57\) found \( \Delta r/r_0 \propto \rho_e \). We see from (8-2) that this requires \( \tau_{\text{in}}(T) \) to be proportional to \( T^{-1} \), independent of \( A \), and proportional to \( \rho_e \). So it's the job of the theorist to cook up an inelastic scattering mechanism with these dependences, right? Not so fast; unfortunately there is a rival theory, which calculates perturbatively corrections to the Ohmic resistance due to electron-electron interactions.\(^{62,63}\) This interaction "theory" also predicts a positive, temperature-dependent incremental resistance in disordered thin metallic wires of the form\(^{63}\)

\[ (\Delta r/r_0)_{ee} = \frac{e^2 v_f}{\pi n A} \left[ \frac{m}{6n} \rho_e \hbar / k_B T \right]^{1/2}. \]  

(8-3)

We see that this approach unambiguously predicts \( (\Delta r/r_0)_{ee} \propto A^{-1} \),
The existence of this rival theory has made the interpretation of experiments on weak localization effects in 1d and 2d systems quite difficult. In 2d, examination of magneto-transport properties has helped to distinguish between the two types of effects, and has shown that both can be important. In 1d such tools have not been successfully applied and the appropriate interpretation of the existing experimental results is still controversial. The fact that the interaction theory does not give the observed dependence of $\Delta r/r_0$ on $\rho_e$ is thus potentially very important; it is the one feature of the present experimental results which can distinguish localization from interaction effects. Therefore, it is essential in calculating $\tau_{in}(T)$ to pay attention to its dependence on $\rho_e$, (or equivalently, on $\ell$, the elastic mean free path), although this was not done previously. The standard calculations of $\tau_{in}(T)$ treat the electrons as plane waves and ignore elastic scattering entirely, so clearly they will be of little use. Instead, it is necessary to consider the inelastic scattering of electrons moving in a random static potential, which are sometimes called diffusons. Schmid, and Abrahams et al. have calculated the diffuson lifetime for inelastic Coulomb scattering; and Schmid and Stone, Joannopoulos and Thouless, have done a similar calculation for inelastic phonon-diffuson scattering. We will present the
results of the phonon-diffusion scattering calculation below; however, neither of these mechanisms gives the correct area and temperature dependence for $\tau_{\text{in}}(T)$ in 1d. The phonon scattering result is nonetheless quite suggestive because it does give the observed dependence of $\tau_{\text{in}}(T)$ on $\rho_e$ correctly, $\tau_{\text{in}}(T) \propto \rho_e$ (as long as $q \ll 1$, where $q$ is the momentum of a thermal phonon). This result appears surprising at first since it says that the inelastic scattering rate decreases as the static disorder increases. However, there are a number of results in the theory of ultrasonic attenuation and dirty superconductors which show that the electron-phonon coupling strength decreases with increasing static disorder. The physics of this phenomenon is still a little puzzling as we shall see below.

There is one inelastic scattering mechanism which is known to give the observed area and temperature dependence for $\tau_{\text{in}}(T)$ in 1d, this is the scattering of electrons from the two-level systems (TLS), or tunneling centers, which are now believed to exist in a wide class of disordered materials. However, the first calculations of $\tau_{\text{in}}$ were done for free electrons, not diffusons. Black et al. found a scattering rate two orders of magnitude too small to agree with experiment, and were, of course, unable to determine the dependence of $\tau_{\text{in}}$ on $\rho_e$. Encouraged by the $\rho_e$-dependence found in the diffuson-phonon calculation, one might hope that a diffuson-TLS scattering calculation would yield the first result for $\tau_{\text{in}}(T)$ which was in complete qualitative agreement with experiment. We perform such a calculation below, for the simplest possible form of the electron-TLS coupling.
First we must digress to introduce and explain the formalism of impurity-averaged perturbation theory which we will employ for the calculation.

No matter how great a distaste one has for Green's functions and diagrammatic summations, it is undeniable that the formalism we describe below has been an absolutely vital tool for the study of transport in disordered metals. To my knowledge, every single quantitative calculation of both weak localization and interaction corrections to the transport coefficients in disordered metals is based on this formalism. And in many cases, it appears to give results in excellent quantitative agreement with experiments in real systems. Thus, it behooves the theorist (and maybe even the experimentalist) to understand how this formalism works. It was invented by Edwards\textsuperscript{78} in 1958 (a good year for localization buffs) and the standard, not-very-helpful reference is section 39.2 of AGD.\textsuperscript{79}

The basis of the formalism can be stated in three sentences. One considers a system of $N$ electrons with coordinates $r_j$ described by a Hamiltonian

$$H = H_0 + H_1 + V_i,$$

where $H_0 = \sum_j p_j^2/2m$, $H_0 + H_1$ is any non-random Hamiltonian you like and $V_i$ is the total potential due to $n$ impurities randomly placed at positions $R_n$ in the system.
Then one treats $H_1 + V_i$ as perturbations and expresses any quantity one hopes to calculate in terms of a set of Feynman diagrams in the standard way. In general, the value of the diagrams will of course depend on the positions of the impurities $R_\alpha$, and so one finds the disorder average of the quantity one is calculating by averaging each diagram over a uniform spatial distribution of the $R_\alpha$. If one takes $H_1$ to be the Coulomb interaction between electrons and one impurity averages the Hartree-Fock diagrams one obtains e.g. "interaction" corrections to the one-particle density of states. If one takes $H_0 + H_1$ to be the Hamiltonian of an electron in a magnetic field and averages conductivity diagrams one obtains "weak localization" corrections to the magnetoresistance and so on. Note, since we are always doing perturbation theory in both $H_1$ and $V_i$ we only expect valid results for disorder which is weak on some scale set by the kinetic energy. The sensible condition which emerges from a careful analysis is $\epsilon T << 1$, thus the energy scale associate with the random potential is the inverse elastic scattering time ($\Xi = 1$).

The simplest case is of course to take $H_1 = 0$ and just do perturbation theory in $V_i$. We illustrate some fundamental properties of this formalism by calculating the one particle Green's function for this case. The fundamental definition of the $T = 0$ one-particle Green's function involves a time-ordered product of Heisenberg field operators averaged in the many-body ground state. However, if we consider a
single particle, instead of a fermi sea of electrons we can use a simpler definition of $G$ which will allow a simple derivation of the Dyson's equation for $G$. This will then be generalized for the fermi sea by means of a modest leap of faith. We define

$$V(r) = \sum_\alpha U(r-R_\alpha) \quad \text{and} \quad G(\omega) = \left[\omega - p^2/2m - V\right]^{-1}.$$  \hspace{1cm} (8-6)

Then, defining $G^0 = [\omega - p^2/2m]^{-1}$, we can formally expand the operator $G$ as

$$G = \left[(\omega - p^2/2m)(I - G^0V)\right]^{-1} = [I - G^0V]^{-1}G^0 = G^0 + G^0VG.$$ \hspace{1cm} (8-7a)

$$G^0 = \sum_{n=0}^{\infty} (VG^0)^n = G^0 + G^0VG.$$ \hspace{1cm} (8-7b)

This is Dyson's equation in operator form. The modest leap of faith (or memory) we now require is that exactly the same equation holds for the one-particle Green's function of a non-interacting system of $N$ electrons in the potential $V_i$, with $G^0$ slightly altered so as to respect the Pauli principle. In the momentum representation

$$G^0(p,\omega) = \frac{1}{\omega - \varepsilon(p) + i\delta \text{sgn}(|p| - k_f)},$$ \hspace{1cm} (8-8)

where $\delta$ is infinitesimal, $\varepsilon(p) = p^2/2m - \varepsilon_f$, and $\text{sgn}(x) = \pm 1$ for $x \gg 0$. The analytic structure of (8-8) ensures that the expectation value of the number operator $a_p^+ a_p$ is one or zero for $p < k_f$ and
respectively. With this substitution we can write (8-7) in the momentum representation as

\[
G(p, p') = G^0(p) \delta_{pp'} + G^0(p) \sum_{p''} V(p-p'') G(p'', p') ,
\]  

(8-9)

where

\[
V(p) = \sum_{r} e^{ip \cdot r} \sum_{\alpha} U(r-R_{\alpha}) = \sum_{\alpha} \sum_{r} e^{ip \cdot r} U(r) \equiv \rho(p) U(p) .
\]  

(8-10)

[We have suppressed \( \omega \) index and replaced integrals by sums for convenience of notation.] The first factor in (8-10) is the impurity density at wave-vector \( p \), the second is the Fourier transform of the impurity potential. Now, iterating (8-9) we get,

\[
G^0(p, p') = G^0(p) \delta_{pp'} + G^0(p) \rho(p-p') U(p-p') G^0(p')
\]

\[
+ G^0(p) \sum_{p''} \rho(p-p'') U(p-p'') G^0(p'') \rho(p''-p') U(p''-p') G^0(p') + ... 
\]  

(8-11)

We can represent (8-11) graphically as

\[
\begin{cases}
G^0(p, p') = \cdots + \frac{V_{pp'}}{p} + \frac{V_{pp'}}{p} + \frac{V_{pp'}}{p} + \cdots \\
\end{cases}
\]  

(8-12)

where the single line is \( G^0(p) \) and the double line is \( G(p, p') \). Now, we are in a position to calculate the impurity-averaged Green's func-
tion, by averaging this series term by term. The random impurity locations now only enter through the density $\rho(p)$. We intend to average $\rho(p)$ over a totally uniform spatial distribution of $R_\alpha$, i.e.

$$P(R_1,...,R_n) = \left(\frac{1}{\Omega}\right)^n,$$

where $\Omega$ is the volume of the system. Since the average system is now translationally invariant, we expect

$$\tilde{G}(p,p') = G(p)\delta_{pp'},$$

to be diagonal in momentum space, and it is easy to check that this automatically falls out of the formalism. To see the results of averaging, note that the first order graph has one factor of $\rho$, and

$$<\rho(p)> = \int \frac{a}{n!} \frac{d^3R_\alpha}{(\Omega)^n} \sum_{\beta=1}^n e^{ip\cdot R_\beta}$$

so

$$\langle \cdots \rangle = n \left[ \sigma^0(p) \right]^2 U(0)\delta_{pp'}, \quad (8-13)$$

The second order graph involves

$$\langle \rho(p)\rho(p') \rangle = \int \frac{a}{n!} \frac{d^3R_\alpha}{(\Omega)^n} \sum_{\beta,\gamma} e^{ip\cdot R_\beta + ip'\cdot R_\gamma}$$

$$= n^2 \delta_{p=0} \delta_{p'=0} + n^2 \delta_{p,-p'}, \quad (8-15)$$

so we get the much more interesting result
The first order graph (8.14) and the first term in (8.16a), simply represent the scattering of the electron from the spatially averaged potential; they contain no interesting physics as we shall see. The second term in (8.16a) is however quite interesting. It represents scattering of the electrons from the same impurity twice, with a total momentum change zero, but with an arbitrary momentum in between the two scattering events. This is very similar to the kind of self-energy correction one gets from the Fock term in the perturbation expansion of the one-particle Green's function of an electron gas interacting through a two-body potential \( [U(r_1 - r_2)]^2 \). Thus, the non-interacting random system behaves in many respects like an interacting non-random system! We will discuss this fascinating phenomenon below, after completing our calculation of the average Green's function. We can represent the result (8.16a) graphically by

\[
\langle - \rightarrow \rightarrow \rightarrow \rangle = n^2 \left| G_0(p) \right|^3 |U(0)|^2 \delta_{pp'}^1 + n \left| G_0(p) \right|^2 \sigma_1(p) \delta_{pp'}^1,
\]

where

\[
\sigma_1(p) = \sum_{p''} |U(p - p'')|^2 G_0(p'').
\]
As we average higher terms in (8-12) we will get a series of diagrams of the form

\[ G(p) = \text{Diagrams} \]

and we can write this in terms of a proper self-energy \( \sigma(p) \) in the usual way

\[
\tilde{G}(p) = \frac{G^0(p)}{1 - \sigma(p)G^0(p)}
\]

Where in general \( \sigma(p) \) consists of all diagrams which cannot be separated into two pieces by cutting a dotted (interaction) line

\[
\sigma(p) = \text{Diagrams}
\]

However, as is usual in perturbation theory, we only calculate \( \sigma(p) \) to lowest order in some small parameter. In this case, we are blessed with two (related) small parameters. If one looks at the way
we averaged products of $p$'s in Eqs. (8-13) and (8-15), it becomes clear that a self-energy diagram with $m$ interaction "dots" will be proportional to $n^m$; so, since we are assuming the impurity density is small, we can exclude all terms (such as the last one in (8-20)) with more than one dot. Second, one can show that any diagram which involves the scattering of an electron from the same impurity more than twice will be smaller by at least a factor of $(\epsilon_F - \epsilon)^{-1}$ than the second term in (8-20). So

$$\sigma(p) = \cdots + \frac{\epsilon}{\epsilon} = nU(0) + n\sigma_1(p) \quad . \quad (8-21)$$

The final simplifying approximation which is conventionally made is to take the impurity potential to be a $\delta$ function in real space, so it is just a constant $U_0$ in momentum space. Then

$$n\sigma_1(p) = nU_0^2 \sum_{p'} C^0(p')$$

$$= \frac{n U_0^2}{(2\pi)^3} \int d^3p' \left[ \omega - \epsilon(p') + i\delta \text{sgn}(|p'| - k_F) \right]^{-1} \quad . \quad (8-22)$$

We change variables in (8-22) to an integral over $\epsilon(p')$ and then extend the lower limit ($\epsilon(p') = -\epsilon_F$) to $-\infty$. This leaves us with the simplest contour integral known to man, $\int_{-\infty}^{\infty} dx [x - \omega \pm i\delta]$, which yields

$$n\sigma_1(p) = i \frac{\text{sgn}(\omega)}{2\pi} \quad , \quad (8-23a)$$
with
\[ \frac{1}{\tau} = \frac{n m k}{\pi} U_0^2 , \quad (8-23b) \]
and thus
\[ \mathcal{G}(p) = \frac{1}{\omega - \epsilon(p) + \frac{1}{2\tau} \text{sgn}(\omega)} , \quad (8-24) \]

where we have absorbed the constant \( nU(0) \) into the definition of \( \epsilon(p) \). This equation is quite interesting, which is why we put so much effort into its derivation. The important thing to notice is that \( \mathcal{G}(p) \) is very much like the Green's function of a system of free electrons interacting through the impurity potential \( U^2 \), and which thus have a finite lifetime for scattering when expressed in the momentum representation, even though \( \mathcal{G} \) is diagonal in that basis. This can never occur for the true Green's function of a system of electrons in a given external potential, it can occur only for the average Green's function. We note further that \( \tau \), as defined in (8-23b), is precisely the elastic scattering time as calculated in the Born approximation. Thus, if we Fourier transform \( \mathcal{G}(p) \) into real space, we find
\[ G(r, r') = \exp \left[ \frac{|r - r'|}{2\xi} \right] , \quad (8-25) \]
which says roughly that the electron scatters elastically after a distance \( \xi \), as we expect. Why does a non-interacting system in a random potential behave on the average like an interacting system
without a potential? If one looks at the derivation of $\bar{G}(p)$ one can get a rough idea of why this is so. Before averaging, the electrons can receive momentum from the impurities, and, whether or not they scatter from the same impurity twice, they can trade momentum back and forth with the lattice. Upon averaging, all scattering sequences which don't conserve momentum will average to zero (for every sequence which changes the electron momentum by $p$, there is an equal and opposite one). And if the sequence involves scattering from different impurities, with uncorrelated positions, then each particular scattering event can involve no exchange of momentum between electrons and lattice. However, if the electrons scatter twice from the same impurity they can first give the lattice momentum and then get that momentum back in a way which conserves momentum for the whole sequence. Thus, by averaging away all scattering sequences which don't conserve total momentum, we can view the electrons as interacting with one another, their interactions mediated by collisions with the same impurity.

However, before one gets too carried away with this description, it is well to note that $\bar{G}(p)$ is not precisely equivalent to the Green's function of a system interacting through $U^2(r)$. The averaging of the Dyson's Eq. (8-9) can never give rise to diagrams with closed fermion loops, and these do, of course, occur for a true interacting system. Such diagrams can be of great physical significance in a true interacting system, as is well known from the RPA approximation to the electron gas. Formally, this feature of
impurity-averaged perturbation theory can be understood by means of a field-theoretic formulation of the localization problem in terms of replicas. However, the physical reason for such a difference is unclear, and one can only wave one's arms so long before getting tired.

There is one further general conceptual feature of this formalism which is not illustrated by the calculation of $\tilde{G}(p)$ and which is extremely important. Often one is interested in using impurity-averaged perturbation theory with $H_1 \neq 0$, either in linear response theory, where $H_1$ is the external field, or in the study of the disordered interacting electron gas which has given rise to the "interaction" theory mentioned earlier. One might think that all that is required is to do standard perturbation theory in $H$, replace all the free electron Green's functions by $\tilde{G}(p)$, and evaluate the resulting diagrams. Not so, valiant reader, there is yet more work to be done. For consider a typical interaction vertex, which involves two Green functions and an external or internal field $A(q)$,

$$
\begin{align*}
\tilde{G}^0(p+q,\epsilon+\omega) & \longrightarrow \quad A(q,\omega) \\
\tilde{G}^0(p,\omega) & \longrightarrow \quad A(q,\omega)
\end{align*}
$$

If one introduces an impurity potential and then replaces each $\tilde{G}^0(p)$ by $\tilde{G}(p)$, one still has not included diagrams of the type shown below
Another way of saying this is that the average two-particle Green's function is not the product of average one-particle Green's functions. Typically, however, one lumps diagrams of the sort shown in (8-26) into a renormalized or corrected interaction vertex.

It can be shown that if one ignores such vertex corrections but uses $\tilde{G}(p)$, that the resulting approximate theory can violate fundamental conservation laws. For example, calculating the conductivity without vertex corrections can in some circumstances lead to an induced current which is not locally conserved. It can also be shown that to be consistent with the self-energy approximation of Eq. (8-20) the appropriately corrected vertex involves summing all "ladder" diagrams of the type shown in the first two terms of (8-26). The physically important result of these considerations is that the effective interaction in the presence of impurities is changed from $A(q,\omega)$ to $A(q,\omega)/(-i\omega \tau + D\tau q^2)$. Thus, the diffusive motion of the electrons strengthens the interactions of the electrons both with themselves and with an external field as $q$ and $\omega$ go to zero. For the case of electrons interacting electromagnetically we can think of the electrons being "slowed down" by the impurities as they scatter from one another and thus interacting more strongly.
More importantly, this renormalized interaction is more divergent as 
$q, \omega \to 0$ which causes the perturbation expansion for many physical quan-
tities to diverge for $d < 2$. It is precisely this divergence, due to 
vertex corrections, of perturbation expansions around the metallic 
state, that indicates that all states are localized in $d < 2$. In 
particular, the logarithmically divergent diagrammatic contributions 
to the conductivity in 2$d$, played an important role in the formulation 
by the Gang of Four of the famous scaling argument described in 
Chapter I.

Now that we understand the basic diffusion formalism; let's apply 
it to calculate the diffusion-TLS inelastic scattering rate in the 
Born approximation. We begin by assuming a simple form for the 
electron-TLS interaction

$$
V_{TLS} = \sum_{m} V_{1}(r - x_{m})\sigma_{m}^{l},
$$

(8-28)

where $r$ is the electronic coordinate, $x_{m}$ are the positions of the TLS, 
and $\sigma_{m}^{l}$ is the Pauli matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ which we define as transferring the 
$m^{th}$ TLS from its ground state to its excited state, or vice-versa.

To be more precise, the state of the system with an electron in state 
$\Psi_{i}(r)$ and the $m^{th}$ TLS, with energy splitting $E$, in its higher or lower 
state will be denoted by $| i, E_{m}^{\pm} \rangle$. Then (8-28) means that

$$
<i, E_{m}^{\pm} | V_{TLS} | j, E_{m}^{\mp} > = \sum_{m} \int \psi_{i}^{*}(r) V_{1}(r - R_{m}) \psi_{j}(r) \equiv < i | V_{\perp} | j > ,
$$

(8-29a)
\begin{equation}
\langle i, E^\pm_m \mid V_{\text{TLS}} \mid j, E^\pm_n \rangle = 0 .
\end{equation}

\(V_1(x-R_m)\) is an unspecified potential the electron sees due to the TLS at \(x_m\). We are not interested in elastic scattering from TLS, so we have not included a diagonal electron-TLS coupling, which would be proportional to \(\sigma^3_m\).

We should state at the outset that in order to get a non-trivial dependence of \(\tau_{\text{TLS}}\) on \(\lambda\), we must assume that the electron-TLS potential has a momentum cut-off at values of \(q\) of order \(1/\xi\). Otherwise, the momentum smearing due to impurities is relatively unimportant and the scattering rate will be that obtained by Black et al. for free electrons, plus corrections of order \((k_f\xi)^{-1}\). This hypothesis, that the TLS is roughly of size \(\lambda\), is intended merely as a provisional hypothesis, to see if the resulting dependence of \(\tau_{\text{TLS}}\) on \(\rho_e\) agrees with experiment. To simplify the calculation we use a trick, first employed by Abrahams et al., called the method of exact eigenstates. We begin by pretending we know the exact eigenstates of the system for a given configuration of impurities, and write down an expression for the scattering rate out of those exact eigenstates due solely to the coupling to the TLS. Those exact eigenstates, designated by \(|i\rangle\) and with energy \(\epsilon_i\), are known as diffuson states. We can use them and the Fermi Golden Rule to immediately write down the average electron-diffuson scattering rate
\[ \bar{R}_{\text{TLS}} = \frac{2\pi}{h} \sum_{m} \left( 1 - f(\varepsilon_j) \right) \left( \left| \langle i \downarrow, E_m^+ | V_{\text{TLS}} | j \downarrow, E_m^- \rangle \right| \right)^2 \delta \left( \varepsilon_i + E_m - \varepsilon_j \right) \frac{e^{-\beta E_m/2}}{2 \cosh(\beta E_m/2)} \]

\[ + \left| \langle i \downarrow, E_m^- | V_{\text{TLS}} | j \uparrow, E_m^+ \rangle \right| \left( \left| \langle i \downarrow, E_m^- | V_{\text{TLS}} | j \uparrow, E_m^+ \rangle \right| \right)^2 \delta \left( \varepsilon_i - E_m - \varepsilon_j \right) \frac{e^{\beta E_m/2}}{2 \cosh(\beta E_m/2)} \] 

(8-30)

Where \( f(\varepsilon) \) is the fermi distribution function, the thermal factors represent the equilibrium probabilities that the TLS are in their ground or excited states, and the bar denotes an impurity average over the exact eigenstates \(|i\rangle\). Using Eqs. (8-29a,b), and changing the sum over TLS states to an integral, we can immediately simplify this to

\[ \bar{R}_{\text{TLS}} = 2\pi \sum_{j} N_T \int_{0}^{\infty} dE \left| V_{\downarrow} | j \uparrow \right| \left( 1 - f(\varepsilon_j) \right) \]

\[ \times \left\{ \delta \left( \varepsilon_i + E - \varepsilon_j \right) B(E) + \delta \left( \varepsilon_i - E - \varepsilon_j \right) B(-E) \right\}, \]

(8-31)

where \( B(E) = \left( e\frac{E}{2} / 2 \cosh \left( \frac{E}{2} \right) \right) \), \( N_T \) is the density of TLS states per unit volume, \( E \) is the energy difference between the TLS ground and excited states, and we have set \( h \) and the system's volume equal to unity. We have written down the scattering rate for an electron in initial state \(|i\rangle\); we actually wish to consider an average scattering rate for all diffusons at the fermi surface (not to be confused with the impurity average). Thus we insert a factor \( \delta(\varepsilon_f - \varepsilon_i) \), divide by the density of electron states at the fermi surface, \( N_0 \), and sum over initial states \(|i\rangle\). The result is further simplified by changing \( E \) to \(-E\) in
the 2nd term of (8-31) and combining terms,

\[ \bar{R}_{\text{TLS}} = \frac{\pi N}{N_0} \int_{-\infty}^{\infty} dE \sech^2 \left( \frac{\beta E}{2} \right) \sum_{ij} |i|V_{ij}|j| |^2 \delta(\epsilon_i + E - \epsilon_j) \delta(\epsilon_i - \epsilon_f) . \]  

(8-32)

Now, to use our impurity-averaging formalism, we need to express the matrix elements \( <i|V_{ij}|j> \) in terms of Green's functions in the momentum basis. We do this using the spectral representation of the advanced and retarded Green's functions

\[ G^\pm(p, p', \omega) = \sum_i \frac{<i|p><i|p'>}{\omega - \epsilon_i \pm i\delta} \]  

(8-33)

so

\[ \frac{1}{2\pi i} (G^- - G^+) = \psi_i(p)\psi_i^*(p') \delta(\omega - \epsilon_i) = \bar{G}(p, p', \omega) . \]  

(8-34)

If we insert four complete sets of momentum states in the expression \( <i|V_{ij}|j><j|V_{ij}|i> \) and use (8-34) we can write

\[ \bar{R}_{\text{TLS}} = \frac{\pi N}{N_0} \int_{-\infty}^{\infty} dE \sech^2 \left( \frac{\beta E}{2} \right) \sum_{P_1, P_2, P_3, P_4} \bar{G}(P_4, P_1, \epsilon_f) \bar{G}(P_2, P_3, \epsilon_f + E) \]  

\[ \times \ V_{\perp}(P_4 - P_3)V_{\perp}(P_2 - P_1) . \]  

(8-35)

Now each of the \( \bar{G} \) can be expanded using Dyson's equation and impurity-averaged in the usual way. However, we know that this will be equivalent, given the standard approximations, to replacing each \( G(p, p') \) by
\( G(p) \) (given by Eq. (8-24)), and summing ladder diagrams to correct the interaction vertex \( V_\perp \). Thus, we can immediately write

\[
\sum_{\{\ell\}} G(p_4,p_1,\varepsilon_f) G(p_2,p_3,\varepsilon_f + E) V_\perp (p_4-p_3) V_\perp (p_2-p_1) = \sum_{q} \Pi(q,E) |V_\perp(q)|^2
\]  

(8-36)

where

\[
\Pi(q,E) = \sum_{\{\ell\}} G(p+q, \varepsilon_f + E) G(p, \varepsilon_f) V_\perp + \sum_{\{\ell\}} G(p+q-p', \varepsilon_f + E) G(p-p', \varepsilon_f) V_\perp + \ldots
\]  

(8-37)

and the solid lines now denote \( \bar{G} = \frac{1}{2\pi i} (\bar{G}^- - \bar{G}^+) \), with \( \bar{G}^\pm = [\omega - \varepsilon(p) \pm i/\tau] \). We shall see that the diagrams in (8-37) form a geometric series which we can sum to arrive at the infamous diffusion pole alluded to earlier. The product \( \bar{G} \bar{G} \) in (8-36) has 4 contributions in principle, \( \bar{G}^+ \bar{G}^+, \bar{G}^- \bar{G}^-, \bar{G}^+ \bar{G}^-, \bar{G}^- \bar{G}^+ \). However, the first two contributions correspond to integrals with poles on the same side of the real axis. This does not mean they are zero, since the integrals only extend from \(-\varepsilon_f\) to \(\infty\), but it does mean they give a contribution down by a factor \((\varepsilon_f \tau)^{-1}\), which can be neglected. The last two contributions are complex conjugates of one another, so we only need to calculate one of them and take twice the real part.
Thus, using the rules discussed earlier, we have

\[
\Pi(q,E) = \left(\frac{1}{4\pi^2}\right)^2 \text{Re}\left\{ \sum_{p} G^+(p+q,\varepsilon_f+E) G^-(p,\varepsilon_f) \right\} 
+ \sum_{p,p'} G^+(p+q,\varepsilon_f+E) G^-(p,\varepsilon_f) nU_0^2 G^+(p+p'+q,\varepsilon_f+E) G^-(p+p',\varepsilon_f) + \cdots \\
= \frac{1}{4\pi^2} \text{Re} \left[ \frac{\Pi_0(q,E)}{1 - nU_0^2 \Pi_0(q,E)} \right]
\]

(8-38)

where

\[
\Pi_0(q,E) = \frac{1}{(2\pi)^3} \int d^3 p \frac{1}{[\varepsilon(|p+q|) - E + i/2\tau][\varepsilon(p) - i/2\tau]}
\]

(8-39)

To evaluate (8-39) we change variables to integrate over \(\varepsilon(p)\), extend the lower limit from \(-\varepsilon_f\) to \(-\infty\), do a simple contour integration, and evaluate the remaining angular integral in the limit \(q,E \to 0\). The result is

\[
\Pi_0(q,E) = \frac{\frac{mk_f}{\pi}}{1 - D\tau q^2 - iE\tau}
\]

(8-40)

Substituting (8-40) into (8-38), using the relation \(1/\tau = \left(\frac{mk_f}{\pi}\right) nU_0^2\), (Eq. (8-23b)), and keeping only lowest order in \(q,E\) gives

\[
\Pi(q,E) = \frac{N_0\tau}{\pi} \left( \frac{D\tau q^2}{(D\tau q^2)^2 + (E\tau)^2} \right)
\]

(8-41)
Using (8-36) and substituting back into (8-35) we get

\[ \bar{R}_{\text{TLS}} = \tau N_T \int_{-\infty}^{\infty} dE \text{sech}^2 \left( \frac{\sqrt{2}E}{2} \right) \int \frac{d^3q}{(2\pi)^3} \frac{D\tau q^2 |V_{\perp}(q)|^2}{[D\tau q^2 + (E)^2]} \]  

Since we are integrating in three dimensions, as long as \( |V_{\perp}(q)|^2 \) is regular as \( q \to 0 \), this integral is well-defined and so we can set \( E_T = 0 \) and perform the energy integral. The result

\[ \bar{R}_{\text{TLS}} \equiv \left( \frac{1}{\tau_{\text{TLS}}} \right) = \frac{12\pi^2}{(eF)^2} \left( \frac{K_B}{M} \right) \left( \frac{N_T}{N_0} \right) (N_0 V_0 V_{\perp})^2 \]  

where \( V_{\perp}^2 = \int_0^1 \frac{dx}{(2\pi)^3} |V_{\perp}(x)|^2 \), \( x = \frac{q}{2k_F} \), and we have written the result in terms of a dimensionless coupling constant \( (N_0 V_0 V_{\perp}) \) which can be estimated from experiment. To the dismay of author and reader alike, this result does not give the experimentally observed linear dependence of \( \tau_{\text{in}} \) on \( \rho_e \). Instead, it predicts \( \tau_{\text{in}} \propto \rho_e^{-1} \), and a scattering rate smaller than that calculated by Black et al. by a factor of about \( (eF)^{-1} \), which is already too small to explain the magnitude of the experimental effect.

This is quite discouraging and on its face appears to rule out an explanation of the thin wire experiments in terms of weak localization effects. However, just because we've turned the crank and this result popped out of the formalism, doesn't mean it contains the right physics, and there is one puzzling feature of this result which requires
explanation. The whole character of the result was changed because of the vertex corrections. A calculation which ignores vertex corrections (which is formally equivalent to replacing $\Pi$ by $\Pi_0$) gives a result in good agreement with experiment. This may be simply fortuitous; however, it is suggestive that for the case of diffuson-phonon scattering, the vertex corrections can be shown to cancel, and the result is that the electron-phonon coupling gets weaker as the disorder increases. There are relatively simple arguments due to Pippard\textsuperscript{71} for why this occurs, and it is unclear why a similar result should not occur for diffuson-TLS scattering. One wonders, therefore, if a better model of the electron-TLS interaction would give results consistent with experiment. Let us now examine the derivation of the diffuson-phonon scattering rate, and see if we can get some insight into why it differs fundamentally from the TLS case.

The most immediate difference is that the standard form of impurity-averaged perturbation theory which we have been using cannot be applied to the electron-phonon problem in a coordinate system fixed with respect to the zero-temperature solid. The reason is that the standard derivation of the electron-phonon interaction takes the electron-ion potential (which depends on the ionic positions $R_\alpha$), assumes the ions undergo a small harmonic displacement from equilibrium $U(R_\alpha^0) = R_\alpha - R_\alpha^0$, and expands the potential around equilibrium to give an interaction term proportional to $U(R_\alpha^0)$. The problem with this is that since the impurities are also embedded in the lattice, their interaction with the electrons must also be expanded in the displacement
U(R₀) which will depend on the random impurity position, Rₐ₁. This has two undesirable features: 1) The total electron-phonon interaction now depends on the random impurity locations, which will create a whole new class of diagrams with impurity interaction lines emanating from the phonon interaction vertex. 2) The interaction with the impurities is no longer static, so the scattering due to impurities is no longer purely elastic.

A clever, but somewhat mysterious method for avoiding these difficulties was first proposed by Blount and worked out by Tsuneto. The idea is to consider the problem in a frame of reference moving with the vibrating lattice. Thus, one defines new electron coordinates \( \mathbf{r} = \mathbf{r}' + U(\mathbf{r}') \), where \( \mathbf{r}' \) is the "lab frame" electron coordinate and \( U(\mathbf{r}') \) is the sinusoidal displacement field which equals \( U(R₀) \) at each lattice site. In this frame, the impurities are fixed so the standard techniques apply, but there is a fictitious force on the electrons, due to the lattice displacements, which comes out of the transformed kinetic energy. This form of the electron-phonon interaction turns out to be

\[
H_{e-p} = \frac{-\hbar^2}{4m} \sum_{b,q,k} \left( (2\mathbf{k}+\mathbf{q}) \cdot \hat{\mathbf{a}}_b(q) \right) \left( (2\mathbf{k}+\mathbf{q}) \cdot \hat{\mathbf{a}}_b(q) \right) \left( \frac{\hbar^2}{2\rho_b(q)} \right) c_{k+q}^+ \left( a_{k-q} + a_{q-k}^+ \right)
\]

(8-44)

where \( \hat{\mathbf{a}}_b(q) \) is the polarization vector of the phonon mode with frequency \( \omega_b(q) \), \( \rho_\mathbf{I} \) is the ionic mass density, \( c^+, c \) are electron creation and annihilation operators and \( a^+, a \) are the analogous phonon operators.
This is an effective kinetic energy of interaction which can be rewritten as an interaction between the lattice distortion $U$ and the electronic stress tensor.\textsuperscript{69} Strictly speaking, the electronic states in (8-44) should be Bloch functions, but we will now approximate them by plane waves, take the $q \to 0$ (low temperature) limit and consider only acoustic modes with a single sound velocity, $\omega(q) = v_s q$.

Again, we write down the Fermi Golden Rule for scattering between impurity eigenstates,

$$
\bar{R}_p = \frac{2\pi}{N_0} \sum_{j \neq i, \phi_j} \left| \langle i, \phi_i | V_{e-p} | j, \phi_j \rangle \right|^2 \delta(E_i - E_j)(1 - f(\varepsilon_j)) \quad ,
$$

(8-45)

where $E_i$ and $E_j$ are the total energies of the initial and final states, $\phi_i, \phi_j$ are many-phonon states. There are two terms which arise out of (8-44) by using (8-44), one corresponding to phonon absorption and the other to emission. For simplicity, we only consider explicitly the case of absorption of longitudinal acoustic phonons. Assuming a thermal distribution of phonons, $n_q = [\exp(\beta \omega_q) - 1]^{-1}$ and performing algebraic manipulations similar to those which led to Eq. (8-35), we reduce (8-45) to

$$
\bar{R}_p = \frac{2\pi}{N_0} \sum_{p_1, p_2, q} \bar{G}(p_2, p_1, \varepsilon_f) \bar{G}(p_1+q, p_2+q, \varepsilon_f + \omega_q) \frac{\Gamma_v(p_1, q) \Gamma_v(p_2, q)}{2 \sinh(\beta \omega_q)}
$$

(8-46a)

where

$$
\Gamma_v(p, q) = -\left( \frac{q}{2m^2 \rho \sqrt{u_s}} \right)^{\frac{3}{2}} (p \cdot \hat{q})^2
$$

(8-46b)
\( \hat{q} \) is a unit vector along \( \vec{q}, \omega_q = v_s \hat{q}, \tilde{G} \) is as defined above (Eq. (8-34)), and \( \lambda = 1 \) again.

Again, we need only consider twice the real part of the \( G^+ G^- \), but now it is convenient to consider the Dyson's equation for the vertex separately. A careful look at our derivation of \( \bar{R}_{TLS} \) will show that we could have reorganized the calculation by defining a bare vertex \( \Gamma_s(q) = V_j(q) \) and solved the Dyson's equation for the corrected vertex \( \tilde{\Gamma}_s(q,E) \),

\[
\tilde{\Gamma}_s(q,E) = \Gamma_s(q) + nu_0^2 \sum \frac{G^+(p+q,\epsilon_f+E)G^-(p,\epsilon_f)}{\mu} \tilde{\Gamma}_s(q,E)
\]

(8-47)

\[
\tilde{\Gamma}_s(q,E) = \frac{\Gamma_s(q)}{1 - n\mu_0^2 \Pi_0(q)} = \frac{\Gamma_s(q)}{iE_\pi + D\tau q^2}
\]

Then we could have calculated \( \bar{R}_{TLS} \) using factorized impurity Green's functions \( \bar{G}^+ \bar{G}^- = \tilde{G}^+ \tilde{G}^- \) and the corrected vertex \( \tilde{\Gamma}_s \), and obtained exactly the same result. Graphically, Eq. (8-47), can be written

\[
\tilde{\Gamma}_s(q,E) = \bigcirc = \bigcirc + \bigcirc = \bigcirc + \bigcirc + \cdots
\]

\[
= \bigcirc \left( 1 + \frac{\omega}{\omega} + \frac{\omega}{\omega} + \cdots \right)
\]

\[
= \frac{\Gamma_s(q)}{1 - \omega} = \frac{\Gamma_s(q)}{d(q,E)}
\]

(8-48a)
is the diffusion pole, the open triangle denotes $\Gamma_s$, and the single lines denotes $\bar{G}$. We have introduced a subscript $s$ to the bare and corrected vertices because (8-48) represents the vertex correction for any interaction which depends only on the momentum transfer, $q$, and not on the external momentum $p$; we shall refer to this as a scalar vertex. The diffusion-phonon interaction vertex (Eq. (8-46)) does depend on the external momentum $p$; we shall refer to this as a vector vertex. Thus, an analysis along the lines of (8-48) does not apply.

To get the right result for the corrected phonon vertex we need to do two things. First, include electronic screening of the bare vertex, to eliminate the unphysical long-range Coulomb interaction which arises from an unscreened longitudinal lattice distortion. We have already done this implicitly by assuming $\omega_q = v_s q$, but for consistency with this assumption, we must introduce RPA-type screening into $\Gamma_v$. For convenience we will resort to diagram algebra a great deal from now on, and leave it to the diligent reader to check that these diagrammatic manipulations correspond to valid manipulations of the mathematical expressions they represent. We denote the bare vertex $\Gamma_v$ by an open triangle, the fully corrected vertex $\tilde{\Gamma}_v$ by a black triangle, and the screened vertex, $\Gamma'_v$, by a cross-hatched triangle. Then $\Gamma'_v$ satisfies the equation
where the wavey line is the bare Coulomb vertex and the ladder denotes that it has been corrected by a ladder sum (cute, eh?). This is the "dirty" RPA approximation. If \( \Gamma_v \) were a scalar vertex we could factor it out of (8-49) and sum immediately to get

\[
\Delta = \Delta + \frac{\Delta}{1 - \Gamma_v} \tag{8-50}
\]

but because of the vector character of \( \Gamma_v \), (8-50) is not correct. Instead, we must iterate (8-49) to next order, then one can factor out a graph depending only \( q \), and sum:

\[
\Delta = \Delta + \frac{\Delta}{1 - \Gamma_v} \rightbracevert \tag{8-51}
\]
The evaluation of the ratio of bubble diagrams appearing in (8-51) is the kind of calculation which is referred to as "straightforward" in AGD; which means that if one hasn't done it before it is likely to be the source of significant personal anguish. Sparing the reader the gory details, it is worth noting for the sake of future generations that it is now essential to include the contribution from $G^+ G^-$, since one is summing over $\omega'$ as well as $p$, and these give a contribution proportional to $N_0$. The result is

$$\Gamma'_v(q,p,\omega) = \Gamma_v(q,p) + gq^2 \left[ \frac{d(q,\omega) - i\omega \tau n U_0^2 \Pi_0(q,\omega)}{d(q,\omega) - i\omega \tau n U_0^2 \Pi_0(q,\omega)} \right], \quad (8-52)$$

where $g = \frac{k_f^2}{3\sqrt{2} \Lambda}\pi$, $d$ is the diffusion pole, and

$$\Pi_2(q,\omega) = \int \frac{d^3 p}{(2\pi)^3} \frac{(\hat{p} \cdot \hat{q})^2}{[(\epsilon(p+q)) - \omega + i/2\tau] (\epsilon(p) - i/2\tau)^2}. \quad (8-53)$$

The evaluation of $\Pi_2(q,\omega)$ is quite similar to that of $\Pi_0(q,\omega)$, and yields

$$n U_0^2 \Pi_2(q,\omega) = 1 - i\omega \tau - \frac{9}{5} D \tau q^2. \quad (8-54)$$

The rest of the calculation is easy. We have only to do the ladder sum for the screened vertex.

The Dyson equation for $\tilde{\Gamma}_v(q,\omega)$ is
\[ \Delta = \Delta + \delta \] \hspace{1cm} (8-55)

Again, we iterate by substituting (8-52) for \( \Gamma_\nu \); the term \( \Gamma_\nu \) remains outside the geometric series, but the remaining terms form a geometric series which sums to give a diffusion pole. The result is

\[
\tilde{\Gamma}_\nu(q, p, \omega) = \Gamma_\nu(q, p) + \frac{g q^{1/2}}{d(q, \omega)} \left[ \frac{d - i \omega \tau n U_\omega^2 \Pi_2}{d - i \omega \tau n U_\omega^2 \Pi_0} \right], \hspace{1cm} (8-56)
\]

and the \( q, \omega \to 0 \) limit of (8-56) is

\[
\tilde{\Gamma}_\nu(q, p, \omega) = \Gamma_\nu(q, p) + g q^{1/2} \left[ \frac{i \omega \tau + 9/5 D \tau q^2}{i \omega \tau + D \tau q^2} \right]. \hspace{1cm} (8-57)
\]

Note, neither term in (8-57) is divergent as \( q, \omega \to 0 \); the diffusion pole has magically disappeared!

Returning finally to (8-46), we can evaluate \( \bar{R}_p \) by factoring the Green's function average and correcting one vertex \( \Gamma_\nu \) using (8-57) [note, correcting both vertices would count the same diagrams twice].

\[
\bar{R}_p = \frac{\text{Re}}{\pi N_0} \left\{ \sum_{p, q} \tilde{G}^-(p, \varepsilon_f) \tilde{G}^+(p+q, \varepsilon_f^+\omega) \frac{\Gamma_\nu(p, q) \tilde{\Gamma}_\nu(p, q, \omega)}{2 \sinh(\beta \omega)} \right\} \hspace{1cm} (8-58)
\]

\[
= \frac{2 \varepsilon_f (\varepsilon_f \tau)}{5 \rho_1 \upsilon_s} \int \frac{d^3 q}{(2\pi)^3} \left[ 1 - \frac{5}{9} \text{Re} \left\{ \frac{i \omega \tau + 9/5 D \tau q^2}{i \omega \tau + D \tau q^2} \right\} \right] \frac{q}{\sinh(\beta \omega)}.
\]
This is an interesting result. The quantity in square brackets is always positive, and only depends weakly on $q$. Since $q$ is cut-off by the thermal factors in (8-58) at values of order $q_T = \frac{K_BT}{v_s}$, in the limit $T \to 0$ the square bracket goes to $4/9$, and all the vertex corrections do is reduce the result obtained using the bare vertex by a factor $4/9$. However, since $\omega_q = (v_s/v_f)(q\xi)$, whereas $(9/5)D\nu q^2 = (3/5)(q\xi)^2$, in a regime of temperature and static disorder where $1 > q_f \xi > \frac{v_s}{v_f} \approx 10^{-2}$, the square bracket is approximately zero, and we have to expand to higher order in $(v_s/v_f)/(q\xi)$. The first non-vanishing correction makes the square bracket proportional to $(v_s/v_f)^2/(q\xi)^2$, which would reduce the temperature-dependence of $R_p$ by two powers of $T$, and totally change its dependence on $\xi$. Although this limit is intriguing and not prima facie unphysical, it's rather difficult to realize. One needs temperature low enough so that $\xi / \lambda$ is not negligibly small (since the resistance rise is proportional to this ratio), but not so low that $q_f \xi \lesssim 10^{-2}$. These conditions do not appear to be met in the thin-wire experiments, where $(q_T)^{-1} \approx 500\text{Å}$ at $5^\circ\text{K}$, and $q_T \xi \approx 10^{-2}$. Thus, the low temperature limit of (8-58) is appropriate; we set the square bracket equal to $4/9$, make the integral dimensionless by rewriting it in terms of $x = \beta v_s q$, and then integrate to find,

\[ \left( \frac{1}{\tau_p} \right)_{3d} = \frac{R_{TLS}}{90} \frac{n^2}{\epsilon_f} \left( \frac{\xi}{v_s} \right) (k_f)^4 \left( \frac{K_B T}{\hbar v_s} \right)^4. \]  

(8-59)

Equation (8-59) is valid for scattering from longitudinal acoustic phonons with $(K_B T/\hbar v_s) \ll 1$ (we have now included electronic scattering
by phonon emission as well, which, not surprisingly, just doubles the scattering rate). The interesting thing about (8-59) is its dependence on \( \lambda \), which gives \( \tau_p \propto \rho_e \). This means inelastic diffusion-phonon scattering does indeed decrease linearly as static disorder increases, exactly the result needed by Giordano et al.\(^{57}\) in order to explain the thin-wire experiments in terms of localization theory. Unfortunately, the temperature-dependence of \( \tau_p \) is much too strong \( (\tau_p \propto T^{-4}) \) to agree with experiment, which finds \( \tau_{in} \propto T^{-1} \).

We have calculated (8-59) assuming three-dimensional phonons, which may appear inconsistent with the assumption that the localization effect is one-dimensional. To quote Bill McMillan, "I may not be too bright, but I ain't dumb".\(^{87}\) The sample is effectively one-dimensional in terms of localization effects if the transverse dimensions are shorter than the localization length (which is \( 10^4 - 10^5 \)Å in the thin wires); whereas it is only one-dimensional in terms of the equilibrium phonon distribution, if the transverse dimensions are shorter than the wave-length of a thermal phonon (so there are very few in thermal equilibrium). For the thin wires this condition is satisfies in the 1°-5°K range, which includes part but not all of the experimental data.\(^{57}\) If we calculate \( \tau_p \) for a 1d phonon distribution all that happens is that in (8-58) the

\[
\int \frac{d^3q}{(2\pi)^3} \to \frac{1}{A} \int \frac{dq_z}{(2\pi)},
\]

which multiplies (8-59) by \( \left( \frac{2}{\pi} \right) (Aq_T)^{-1} \), giving

\[
\left( \frac{1}{\tau_p} \right)_{1d} = \frac{\pi}{45} \left( \frac{\epsilon_f}{\rho_f V_s} \right) (k_f \ell) \frac{1}{A} \left( \frac{K T}{N V_s} \right)^2.
\]

(8-60)
This weaker $T$-dependence is closer to the experimental result; but unfortunately, the additional area-dependence of $\tau_p$ will change the prediction of (8-2) for the resistance rise to $(\Delta r/r_0) \propto A^{-1/2}$, which is not seen experimentally. So, although $\tau_p$ has the anomalous dependence on $\rho_e$ suggested by the thin-wire experiments, it is not a very promising candidate for the dominant inelastic scattering mechanism in the experiments.

So where does this leave us in terms of our inelastic TLS scattering result of Eq. (8-43)? Believe it or not, we've actually gained some physical insight into why $\tau_p$ depends anomalously on $\rho_e$ and $\tau_{TLS}$ does not. The difference between the scalar vertex in the electron-TLS interaction and the vector vertex in the electron-phonon interaction turns out to be somewhat of a red herring. The physically significant difference between the two cases appears to be that phonons have an energy-momentum dispersion relation and TLS presumably do not. If the TLS had such a dispersion relation, then instead of summing separately over momentum transfer, $q$, and TLS energy, $E$, as we did in (8-42), we would have to sum only once simultaneously over $q$ and $\omega_q$, as we do in the phonon case (Eq. (8-58)). This would give

$$
\bar{R}_{TLS} \propto \ell \sum \frac{|v_L(q)|^2 D \tau_0}{q (\omega_q \tau)^2 + (D \tau q)^2} \left( \text{sech}^2(\beta \omega_q) \right) . \tag{8-61}
$$

In the $T \to 0$ limit, the $(D \tau q)^2$ term would be negligible compared to $(\omega_q \tau)^2$, the factors of $\ell^2$ would cancel in the integral, and the
scattering rate would be proportional to $\lambda$, just as in the phonon scattering case.

Thus, to summarize, neither the diffuson-TLS or the diffuson-phonon inelastic scattering rates give results in full agreement with Giordano's thin-wire experiments. Phonons give the correct (anomalous) dependence of $\tau_{in}$ on $\rho_e$, but their momentum-energy dispersion relation implies a rapid variation of the density of states with energy which gives too rapid a temperature-dependence for $\tau_{in}$. TLS, on the other hand, have a flat density of states and give the correct T-dependence; but the very fact that they have no energy-momentum dispersion relation means that vertex corrections will remove the anomalous dependence of $\tau_{in}$ on $\rho_e$. Thus, the experimental results on thin wires remain somewhat of a puzzle from the point of view of localization theory. We shall discuss this further in our concluding remarks.
Chapter IX. Conclusions

If one tries now to assess the state of the theory of one-dimensional localization, two conclusions seem to be very clear: 1) Our understanding of the zero-temperature 1d localization problem is virtually complete. 2) Our understanding of the role of localization effects in real, interacting, finite-temperature quasi-one-dimensional systems is far from complete.

At zero temperature in 1d, we have shown: 1) The Landauer formula provides a convenient, rigorous expression for the resistance, of great computational utility. 2) The unusual statistical behavior of the resistance as a function of sample length can be understood and predicted using the properties of lognormal distributions. 3) The random phase model is generally correct and appears to contain the essential physics of the localization phenomenon.

At finite temperature we have shown: 1) The fact that the localization length is a well-behaved random variable implies that the zero-temperature statistical fluctuations discussed above will not affect experimental results in the weakly or intermediately localized regimes; although they may be significant in the exponential hopping regime. 2) Explicit calculations of $\tau_p$ and $\tau_{TLS}$ do not yield inelastic scattering rates in full quantitative or qualitative agreement with experiment. Thus, whether the thin-wire experiments are measuring localization effects or interaction effects or both is still unclear.

It is interesting to speculate on what interpretation of these
experiments seems most likely, given what we know at the moment. A very recent experiment by White et al. is quite interesting in this regard. All previous experiments had been done on extremely dirty systems to maximize the size of the effect; however White et al. studied relatively clean copper wires and found results in good agreement with the interaction prediction. The reason that this is interesting is that these systems fall clearly in the range of validity \((k_f \ell \gg 1)\) of the impurity-averaged perturbation theory used to calculate both the inelastic diffusion length and the interaction correction to the resistance. Most of the other experiments are on systems with \(k_f \ell \) approaching unity, so the use of perturbation theory is becoming suspect. The fact that the interaction theory naturally predicts the correct experimental \(T\)-dependence and gives good quantitative agreement for the effect in copper would appear to make it the front-runner.

There is still the unresolved problem of the anomalous \(\rho_e\)-dependence of the effect in Giordano's experiments. It does not appear from our calculation as if any minor tinkering with the electron-TLS interaction will give \(\tau_{TLS}\) the right \(\rho_e\)-dependence. It is important to remember that if the TLS turn out to be truly localized objects on the scale of \(\ell\), then \(\tau_{TLS}\) can have no interesting dependence on \(\rho_e\) at all, since scattering is then dominated by momentum transfers greater than \(\ell^{-1}\). One should not, however, rule out the possibility that a deeper understanding of the TLS will allow a scattering behavior consistent with experiment, in particular there is the vague feeling that at
low temperatures the electron may somehow exchange energy with the TLS, due to their flat density of states, but exchange momentum with acoustic phonons. But, it should also be pointed out that it is very difficult to take a single material and vary \( \rho_e \) while keeping all other relevant parameters constant, and estimates of \( \rho_e \)-dependence based on combining data from several materials fabricated differently are suspect. Thus, one has to worry about the reliability of present experimental results, and further corroboration of the anomalous \( \rho_e \)-dependence would be welcome.

Another interesting possibility for distinguishing experimentally between localization and interaction effects is to try to observe the cross-over from the weakly localized to the intermediately localized regimes. Localization theory predicts that the inverse power-law rise of the resistance should double when the inelastic diffusion length becomes greater than the localization length in 1d. This regime is almost within the reach of present experiments, and its observation would certainly strengthen our belief in the correctness of Thouless' physical picture of the finite temperature effects of localization in one dimension.
References


25. A recent article by B. Shapiro, Phys. Rev. Lett. 48, 823 (1982), uses a real-space renormalization approach based on a one-dimensional scaling law, to predict the properties of the Anderson transition in any dimension. I believe this result provides strong support for my claim that the one-dimensional case is indeed of more general physical significance.
27. This is just a convenient device for eliminating surface contributions to the integral, which we expect to be absent on physical grounds.
28. The results of Langreth and Abrahams (Ref. 26) are not simple even in the case of many channels with the same velocity; and Langreth has argued (Thouless, private communication) that this result is correct for several communicating channels in space, although not for several communicating spin channels, the case I have treated. So apparently there are even further subtleties lurking in this innocent-looking problem.


37. They show that wave-function amplitudes on even sites decays as \(\exp[-x^2/\lambda]\) if those on odd sites grow as \(\exp[+x^2/\lambda]\), or vice-versa, they then ignore the exponentially growing piece.


40. The fact that each transfer matrix is described by exactly three independent parameters, in general, follows from the unitarity of the S-matrix. To be consistent with previous notation we have denoted R/T for the whole chain as $\rho_N$ and for a single scatterer by $z_N$; Lambert and Thorpe do exactly the opposite.


42. This statistical test is based on the following observations. The quantities

$$\sqrt{2}C_q = \frac{\sqrt{2}}{m} \sum_{i=1}^{m} \cos(q\phi_i)$$

and

$$\sqrt{2}S_q = \frac{\sqrt{2}}{m} \sum_{i=1}^{m} \sin(q\phi_i)$$

have a joint normal distribution with zero mean and unit variance in the limit $m \to \infty$. For $q = 1$ this is familiar since $C_q$ and $S_q$ are the x and y components of a random walk in two dimensions; and it is easy to prove it for general $q$ by looking at the characteristic function of $C_q$ and $S_q$ as $m \to \infty$. Then the quantity

$$M_q = 2(C_q^2 + S_q^2)$$

is a chi-squared variable with 2 degrees of freedom, and $T = \sum_{q} M_q$ is chi-squared distributed with 2q degrees of freedom.


45. One such exceptional case is the band center state for purely off-diagonal disorder discussed in Chapter IV. Other exceptions appear to exist in the disordered Kronig-Penny model and are discussed by M. Ya. Azbel, Solid State Comm. 37, 789 (1981).


51. This distance is often called the Thouless length by localization theorists, since its importance in this context was first pointed out by Thouless (see Ref. 5).


60. It may be worth reminding the reader why this relation holds. It is clear that the value of $\ln \rho$ at a given energy should be closely related to the inverse localization length $\gamma(E)$, which gives the decay rate of the solution of the Schrodinger equation at that energy. For example, in the Anderson model with $N$ sites $\gamma(E)$ may be defined by fixing the wave-function amplitudes at one end of the chain $a_1 = 1$ and $a_0 = 0$; then $\gamma(E) = \frac{1}{N-1} \ln |a_N|$. We have shown in Chapter III that given the Anderson Hamiltonian and the wave-function amplitudes on two adjacent sites all the others may be determined recursively by the equation

\[
\begin{pmatrix}
a_N \\
a_{N-1}
\end{pmatrix} = \begin{pmatrix}
\frac{E-\epsilon_i}{V_{i,i+1}} - \frac{V_{i-1,i}}{V_{i,i+1}} \\
1
\end{pmatrix}
\begin{pmatrix}
a_{i+1} \\
a_i
\end{pmatrix}
\]

where the $\epsilon_i$ are the site energies and the $V_{ij}$ are the nearest neighbor couplings, both of which may be random. Taking $a_1 = 1$ $a_0 = 0$ we see that $a_N = \frac{P^{11}}{N}$ and $\gamma = \frac{1}{N-1} \ln \left| \frac{P^{11}}{N} \right|$. On the other hand, we have shown (Eq. (3-13)) that $R/T$ in general is given by

\[
\frac{1}{N} \ln (R/T) = 2 \gamma,
\]

where $P^{ij}$ are the nearest neighbor couplings, both of which may be random. Taking $a_1 = 1$ $a_0 = 0$ we see that $a_N = \frac{P^{11}}{N}$ and $\gamma = \frac{1}{N-1} \ln \left| \frac{P^{11}}{N} \right|$. On the other hand, we have shown (Eq. (3-13)) that $R/T$ in general is given by

\[
\sum_{ijkl} C_{ijkl} P^{ij}_{N} P^{kl}_{N} (where the $C_{ijkl}$ are constants of order unity, except very near the band edge) and that the $P^{ij}_{N}$ are all the same except for boundary effects. Thus $\frac{1}{N} \ln (R/T) = \frac{2}{N} \ln |P^{11}_{N}| = 2 \gamma$, plus corrections of order $1/N$. A similar argument can be made for any disordered one-dimensional model, since the wave-function amplitudes on either side of the disordered region can always
be related by a transfer matrix, which typically will cause that amplitude to grow by a factor \( C \sqrt{R/T} \), where \( C \) is a constant of order unity. It is not trivial to show that this definition of \( \gamma \), in terms of a growth rate of the wave-function amplitude given fixed boundary conditions at one end of the system, is essentially equivalent to a more natural definition in terms of the decay rate of an eigenfunction with amplitude fixed to be equal at both ends of the chain. Thouless, (J. Phys. L5, 77 (1972)), has shown that both definitions lead to the same expression for \( \gamma \) in terms of the density of states. This touches upon a well-known subtlety in the theory of one-dimensional localization, see R. E. Borland, Proc. Roy. Soc. A274, 529 (1963) and for a recent discussion, J. C. Kimball, Phys. Rev. B 24, 2964 (1981). For a wire of finite cross-section one must define the localization length somewhat differently and the result \( \gamma = \frac{1}{2} \ln(R/T) \) may have corrections, but one expects these to be small if \( m \), the number of scattering channels is much less than \( N \).

61. It may appear inconsistent to ignore intrinsic fluctuations in \( \hat{l}_1(T) = \sqrt{\hat{v}_f \tau_{\text{in}}} \), since it depends on \( \hat{l} \), the mean free path, and in 1d the mean free path is just half the localization length, which we are treating as the underlying random variable. Actually, the correct statement (see Ref. 52) is that the mean free path is proportional to the average localization length, and is not therefore a random variable itself. Nonetheless, it is
not clear to me on physical grounds that the inelastic length scale (as opposed to the inelastic time scale) should not be treated as a random variable. A correct finite temperature theory will presumably show exactly how the resistance depends on the random length scale set by the microscopic disorder and thus answer this question. However, the assumption that $l_i$ is non-random seems adequate for the purposes of the present argument.


70. A. D. Stone, J. D. Joannopoulos and D. J. Thouless, unpublished.


83. R. Schrieffer, Superconductivity, Benjamin, New York (1904), Chapter 8.
84. We are entitled to integrate momentum transfers over three-dimensional q-space as long as the transverse dimensions of our system are much greater than $\lambda$.
Figure Captions

Fig. 4-1: Plot of $\langle \ln(1+p) \rangle$ vs. $N$ for a 1d Anderson Model with purely off-diagonal disorder for several non-zero values of energy. The results are for $W_V = 0.25$ and an ensemble of 500 chains. The insert plots $N$ times the relative variance

$$\Gamma = \frac{\langle (\ln(1+p) - \langle \ln(1+p) \rangle)^2 \rangle}{\langle \ln(1+p) \rangle^2}$$

vs. $N$. Averages are over an ensemble of 500 chains.

Fig. 4-2: Plot of $\langle \ln(1+p) \rangle$ vs. $\sqrt{N}$ for purely off-diagonal disorder and $E = 0$ for $W_V = 0.25$ and an ensemble of 500 systems. The solid line is the analytic result of Ch. IV which is only exact an $N \to \infty$.

Fig. 4-3: Plot of $\langle \ln(1+p) \rangle$ vs. $N$ for both diagonal and off-diagonal disorder at $E = 0$. Results are for an ensemble of 400 chains with $W_V = 0.25$ and several values of $W_\varepsilon$.

Fig. 4-4: Plot of $\ln[\langle 1+p \rangle]$ vs. $N$ for both diagonal and off-diagonal disorder at $E = 0$. The solid curve is the analytic result of Eq. (3-22). The squares are the numerically computed average. Averages were taken over an ensemble of 1000 chains.
Fig. 5-1a-1b: Histograms of \( P(v) \) in arbitrary units for an ensemble of 5000 chains. In Fig. 1a chains are of length 1000 atoms, \( W = 0.25 \), and \( E = 1.0 \). \( P(v) \) is uniform with 95% confidence, as described in text. In Fig. 1b, \( E = 0 \), \( W = 0.25 \), and chains are of length 10,000 atoms. This histogram fails a \( \chi^2 \) test for uniformity, as did every ensemble at \( E = 0 \).

Fig. 5-2a-2b: Histograms of \( P(v) \) illustrating the approach to uniformity and its independence of \( W \). Figs. b and d pass \( \chi^2 \) test for uniformity whereas Figs. a and c fail. Note the similarity of Figs. a to c and b to d, even though the disorder has been varied by a factor of 100. For these parameters, uniformity is reached around \( N = 30 \). Ensemble size is 5000.

Fig. 5-3a-3b: Histograms illustrating the dependence of \( P(v) \) on the product \( NE \) which is held constant here. Ensemble size is 5000.

Fig. 5-4: Histogram of \( P(v) \) illustrating large disorder behavior. The peak is precisely at \( \pi + 2\cos^{-1}(E/2) = 5\pi/3 \) as predicted. The shape of \( P(v) \) changes very little with \( N \) as expected. Ensemble size is 1000.
Fig. 6-la: Plot of $\ln(\bar{p})$ vs. number of atoms for Anderson model with purely diagonal disorder of half-width $W/V = 1/\sqrt{2}$. Circles are the numerically computed average over 1000 systems; the thick curve is calculated from Eq. (6-19a). The values of $\sigma$ and $\gamma$ are obtained numerically. The thin solid line is the analytic result, $\ln[<p>]$ from Chapter III.

Fig. 6-1b: Plot of $\ln(\bar{g})$ vs. number of atoms for the same system parameters as Fig. 6-la. Again, the solid curve is the analytic result of (6-19b) and the circles the numerical values.

Fig. 6-2: Plot showing the fit of Eq. (6-19a) to the data of Fig. 4-4. Here the dashed line is the analytic result for $\ln\bar{p}$, the squares are numerical values of $\ln\bar{p}$, and the solid line is $<\ln(1+\rho)>$. 
\[ \langle \ln (1+\rho) \rangle \]

- \( E = 0 \)
- \( W_\epsilon = 0 \)
- \( W_\nu = 0.25 \)
The graph shows the behavior of $\langle \ln(1+\rho) \rangle$ as a function of $N$ for different values of $W_\epsilon$. The lines correspond to different values of $W_\epsilon$: $E = 0$, $W_\epsilon = 0.25$, $W_\epsilon = 0.01$, $W_\epsilon = 0.1$, and $W_\epsilon = 1.0$.
\[ \ln \langle (1 + p) \rangle \]

E = 0

\[ W_\epsilon = 0.1 \]

\[ W_v = 0.25 \]
Fig. 5-1

a) $E = 1.0$
$N = 1000$
$W = 0.25$

---

b) $E = 0$
$N = 10,000$
$W = 0.25$
Fig 5-3

(a) $E = 0.05$
$N = 50$
$W = 0.1$

(b) $E = 0.25$
$N = 10$
$W = 0.1$
$E = 1.0$

$N = 2$

$W = 50.0$

Fig. 5.4
E = 0
$W_\epsilon = 0.1$
$W_v = 0.25$