

The coefficient of viscosity is therefore the zero-field value $\eta(0)$, with $S_1(\zeta)$ replacing the former $S(\zeta)$:

$$\eta(H) = \eta(0)[1 + 4\omega_c^2\tau^2]^{-1}. \quad (8.7)$$

The magnetic field suppresses the viscosity by shortening the mean free path in the direction of transport. Apart from differences in the magnitude of τ , the term $4\omega_c^2\tau^2$ in (8.7) replaces $\omega_c^2\tau^2$ in the analogous result for the conduction problem. This is owing to charge transport being reversed by turning through 180° while transverse momentum transport is reversed by turning through 90° , or in one-half the time. The assumption of

a time of relaxation limits the validity of (8.7) to $T \gg \Theta$. However, as shown by Sondheimer and Wilson for the electrical conductivity,¹¹ such a formula is probably more widely applicable than its derivation would suggest.

ACKNOWLEDGMENT

The author wishes to thank Professor Lars Onsager, under whose guidance this research was undertaken, for his interest and advice during the course of the work.

¹¹ E. H. Sondheimer and A. H. Wilson, Proc. Roy. Soc. (London) **190**, 435 (1947).

Absence of Diffusion in Certain Random Lattices

P. W. ANDERSON

Bell Telephone Laboratories, Murray Hill, New Jersey

(Received October 10, 1957)

This paper presents a simple model for such processes as spin diffusion or conduction in the "impurity band." These processes involve transport in a lattice which is in some sense random, and in them diffusion is expected to take place via quantum jumps between localized sites. In this simple model the essential randomness is introduced by requiring the energy to vary randomly from site to site. It is shown that at low enough densities no diffusion at all can take place, and the criteria for transport to occur are given.

I. INTRODUCTION

A NUMBER of physical phenomena seem to involve quantum-mechanical motion, without any particular thermal activation, among sites at which the mobile entities (spins or electrons, for example) may be localized. The clearest case is that of spin diffusion^{1,2}; another might be the so-called impurity band conduction at low concentrations of impurities. In such situations we suspect that transport occurs not by motion of free carriers (or spin waves), scattered as they move through a medium, but in some sense by quantum-mechanical jumps of the mobile entities from site to site. A second common feature of these phenomena is randomness: random spacings of impurities, random interactions with the "atmosphere" of other impurities, random arrangements of electronic or nuclear spins, etc.

Our eventual purpose in this work will be to lay the foundation for a quantum-mechanical theory of transport problems of this type. Therefore, we must start with simple theoretical models rather than with the complicated experimental situations on spin diffusion or impurity conduction. In this paper, in fact, we attempt only to construct, for such a system, the simplest model we can think of which still has some expectation of representing a real physical situation

reasonably well, and to prove a theorem about the model. The theorem is that at sufficiently low densities, transport does not take place; the exact wave functions are localized in a small region of space. We also obtain a fairly good estimate of the critical density at which the theorem fails. An additional criterion is that the forces be of sufficiently short range—actually, falling off as $r \rightarrow \infty$ faster than $1/r^3$ —and we derive a rough estimate of the rate of transport in the $V \propto 1/r^3$ case.

Such a theorem is of interest for a number of reasons: first, because it may apply directly to spin diffusion among donor electrons in Si, a situation in which Feher³ has shown experimentally that spin diffusion is negligible; second, and probably more important, as an example of a real physical system with an infinite number of degrees of freedom, having no obvious oversimplification, in which the approach to equilibrium is simply impossible; and third, as the irreducible minimum from which a theory of this kind of transport, if it exists, must start. In particular, it re-emphasizes the caution with which we must treat ideas such as "the thermodynamic system of spin interactions" when there is no obvious contact with a real external heat bath.

The simplified theoretical model we use is meant to represent reasonably well one kind of experimental situation: namely, spin diffusion under conditions of

¹ N. Bloembergen, *Physica* **15**, 386 (1949).

² A. M. Portis, *Phys. Rev.* **104**, 584 (1956).

³ G. Feher (private communication).

“inhomogeneous broadening.”⁴ We assume that we have sites j distributed in some way, regularly or randomly, in three-dimensional space; the array of sites we call the “lattice.” We then assume we have entities occupying these sites. They may be spins or electrons or perhaps other particles, but let us call them spins here for brevity. If a spin occupies site j it has energy E_j which (and this is vital) is a stochastic variable distributed over a band of energies completely randomly, with a probability distribution $P(E)dE$ which can be characterized by a width W . Finally, we assume that between the sites we have an interaction matrix element $V_{jk}(r_{jk})$, which transfers the spins from one site to the next. V_{jk} may or may not itself be a stochastic variable with a probability distribution. If one thinks of the mobile entities as up or down electron spins which can occupy various impurity sites, such as color or donor centers, in a crystal, then the random energies E_j are the hyperfine interactions with the surrounding nuclei—Si²⁹ for the donors, alkali and halide nuclei for color centers—and $P(E)$ is the line-shape function. In this example, V_{jk} is that part of the interaction which allows an up spin on atom j to flip down while a down spin on k flips up, and the simple process we study is the motion of a single “up” spin among “down” spins.

The “impurity-band” example would again make the sites donors or acceptors, but the E_j 's would be energy fluctuations of the donor ground state caused perhaps by Coulomb interactions with randomly placed charged centers; the moving entity would be a single ionized donor. We would have to assume the states of the different donors to be orthogonal, which is no restriction if V is arbitrary. More generally, the situation described by our theorem probably holds in the low-concentration limit and the low-energy tail of almost any model for an impurity band.

One important feature which is missing from our simple model is contact with any external thermal reservoir. When the present theorem holds, some such contact will actually control the transport processes. Our purpose is only to show that the model in itself provides no such reservoir and permits no transport, in spite of its large size and random character; study of the real relaxation and transport processes must come later.

Our basic technique is to place a single “spin” on site n at an initial time $t=0$, and to study the behavior of the wave function thereafter as a function of time. Our fundamental theorem may be restated as: if $V(r_{jk})$ falls off at large distances faster than $1/r^3$, and if the average value of V is less than a certain critical V_c of the order of magnitude of W ; then there is actually no transport at all, in the sense that even as $t \rightarrow \infty$ the amplitude of the wave function around site n falls off

rapidly with distance, the amplitude on site n itself remaining finite.

One can understand this as being caused by the failure of the energies of neighboring sites to match sufficiently well for V_{jk} to cause real transport. Instead, it causes virtual transitions which spread the state, initially localized at site n , over a larger region of the lattice, without destroying its localized character. More distant sites are not important because the probability of finding one with the right energy increases much more slowly with distance than the interaction decreases.

This theorem leaves two regions of failure (and therefore transport) to be investigated, namely, $V \propto 1/r^3$, as in spin diffusion by dipolar interactions, and $V \sim W$, or the high-concentration limit. In both cases, the methods used to prove the fundamental “nontransport” theorem will probably allow us to outline an approach to the transport problem. In the $1/r^3$ case, we show that transport may be much slower than the estimates of reference 2 would predict. In the case $V \sim W$, we show that transport finally occurs not by single real jumps from one site to another but by the multiplicity of very long paths involving multiple virtual jumps from site to site.

II. SUMMARY OF THE REASONING

Since the mathematical development is fairly complicated and involves lengthy consideration of each of a number of points, we should like to summarize the reasoning rather fully in this section, leaving the proofs and details to later sections. First, then, let us set up the simple model which we study. The equation for the time-dependence of the probability amplitude a_j that a particle is on the site j is:

$$i\dot{a}_j = E_j a_j + \sum_{k \neq j} V_{jk} a_k. \quad (1)$$

Here we measure energies in frequency units, so we can set $\hbar=1$. Equation (1) simply restates the assumptions about the model made in the Introduction.

We study the Laplace transform of the equation (1): let

$$f_j(s) = \int_0^\infty e^{-st} a_j(t) dt, \quad (2)$$

and then

$$i[sf_j(s) - a_j(0)] = E_j f_j + \sum_{k \neq j} V_{jk} f_k. \quad (3)$$

The variable s must be studied as an arbitrary complex variable with positive or zero real part.

The transport problem which interests us is: suppose we know the probability distribution $|a_j(0)|^2$ at time $t=0$, and that it is appreciable only in a certain range of frequency E_j or space r_j . Then how fast,⁵ if at all, does this probability distribution diffuse away from this region?

⁴ A. M. Portis, Phys. Rev. 91, 1071 (1953).

The simplest question we can ask is to assume $a_0(0)=1$ for a particular atom $j=0$ and inquire how a_j varies with time, or $f_j(s)$ with s . In particular, for very small real part of s we are studying the behavior as $t \rightarrow \infty$; for instance, $\lim_{s \rightarrow 0^+} s f_j(s)$ is in fact $\langle a_j(\infty) \rangle_{av}$.

Equation (3) can be written

$$f_j(s) = \frac{i\delta_{0j}}{is - E_j} + \sum_{k \neq j} \frac{1}{is - E_k} V_{jk} f_k(s). \quad (4)$$

In one approach to ordinary transport theory⁵ this equation is solved by iteration, in which the equations (4) not involving $f_0(s)$ are solved for $f_j(s)$ in terms of $f_0(s)$:

$$f_j(s) = \frac{1}{is - E_j} V_{j0} f_0(s) + \sum_k \frac{1}{is - E_j} V_{jk} \frac{1}{is - E_k} V_{k0} f_0(s) + \dots \quad (5)$$

Then the zeroth equation becomes

$$f_0(s) = \frac{i}{is - E_0} + \sum_k \frac{1}{is - E_0} V_{0k} \left(\frac{V_{k0}}{is - E_k} + \sum_l \frac{1}{is - E_k} V_{kl} \frac{1}{is - E_l} V_{l0} + \dots \right) f_0(s). \quad (6)$$

Let us call the quantity

$$\sum_k \frac{(V_{0k})^2}{is - E_k} + \sum_{k,l} \frac{V_{0k} V_{kl} V_{l0}}{(is - E_k)(is - E_l)} + \dots = V_c(0). \quad (7)$$

In many cases the first term suffices. Studying this first term, we see that it can be written

$$V_c(0) = \sum_k (V_{0k})^2 \left(\frac{-E_k}{s^2 + E_k^2} - \frac{is}{s^2 + E_k^2} \right). \quad (8)$$

In the limit as $s \rightarrow 0$, the first part of this is obviously just the second-order perturbation $-\Delta E^{(2)}$ of the energy. The second part may be written

$$\lim_{s \rightarrow 0} (V_c) = -i \sum_k (V_{0k})^2 \delta(E_k) - is \sum_{k, E_k \neq 0} \frac{(V_{0k})^2}{E_k^2} \equiv -\frac{i}{\tau} - isK. \quad (9)$$

Of course, if the first of these terms, the usual transition probability, is finite the second is indeterminate. We include both in order to see what happens if the first term does vanish.

⁵ See, for example, W. Kohn and J. M. Luttinger, Phys. Rev. **108**, 590 (1957).

Now the equation for $f_0(s)$ is

$$f_0(s) = \frac{i}{is - E_0} + \frac{1}{is - E_0} \left(-\Delta E^{(2)} - \frac{i}{\tau} - isK \right) f_0(s).$$

The solution for f_0 is

$$f_0 = \frac{i}{is(1+K) + (i/\tau) - (E_0 - \Delta E^{(2)})}. \quad (10)$$

If τ is finite, one gets the usual result of perturbation theory:

$$f_0(s) = \frac{1}{s + (1/\tau) + i(E_0 - \Delta E^{(2)})}, \quad (10A)$$

which represents a state of perturbed energy $E_0 - \Delta E^{(2)}$ decaying at a rate $e^{-t/\tau}$. If, on the other hand, τ is infinite [$\text{Im}(V_c) \rightarrow 0$ as $s \rightarrow 0$], then the constant K enters and the amplitude is

$$f_0(s) = \frac{1}{s(1+K) + i(E_0 - \Delta E^{(2)})}. \quad (10B)$$

This is a state of the same perturbed energy which does not decay, but has a finite amplitude a_0 ($t \rightarrow \infty$) reduced from unity by the ratio $1/(1+K)$. K is then simply a measure of how much this state has spread to the neighbors by virtual transitions, and has nothing to do with real transport.

Our technique, then, will be to study the behavior of the quantity $V_c(s)$ defined in (7) by an infinite series, just as in the usual transport theory. However, the quantities entering into (7) are completely different in character because of the fact that we have chosen to start with localized states of random energies E_j . In the usual case, there are an infinite number of states j connected to any state 0 by infinitesimal matrix elements V_{0j} , so that within any small range of energies there will be many possible energy-conserving transitions, no one of which takes place with particularly large probability. In such a case the first term of (9) is a meaningful limit of a certain integral. Here, only a few V_{0j} 's are large, and the energies they lead to are stochastically distributed, so that whether or not energy can be conserved is a probability question.

We find that the quantity $V_c(s)$ must be studied as a probability variable: that is, we pick a starting atom 0 and an arbitrary energy E [imaginary part of s ; $\text{Re}(s) \rightarrow 0$] and study the probability distribution of V_c . Our study then resolves itself into three parts: First, we study the first term (8); second, we discuss the convergence of the series of higher order perturbations. Both of these questions we can resolve in the sense that there is a region in which, with probability unity, $\text{Im}(V_c) \rightarrow 0$ as $\text{Re}(s) \rightarrow 0$ and the series is convergent. These two parts we shall discuss here briefly, and expand upon in Secs. III and IV. Finally, we must decide whether this kind of convergence in a probability

sense is meaningful, and in particular whether the choice of an arbitrary energy is correct. Since this seems reasonable from the start, we shall not go into it further here, but reserve the discussion for Sec. V. We find there that this convergence means that the states are localized, but that it is not easy to assign a correspondence between the perturbed and unperturbed states.

Let us, then, go ahead with the first two questions. The important quantity for the first is

$$\text{Im}(V_c) = -s \sum_k \frac{|V_{0k}|^2}{s^2 + E_k^2}.$$

Let

$$\sum_k \frac{|V_{0k}|^2}{s^2 + E_k^2} \equiv X(s). \quad (11)$$

We note that X also represents the quantity

$$X(s) = \sum_{j \neq 0} \frac{|f_j(s)|^2}{|f_0(s)|^2} \quad (11')$$

in first order, as is clear from (5). This points up the interpretation that as $s \rightarrow 0$ a finite X means no real transport.

Using the Holtmark-Markoff⁶ method, in the next section we calculate the probability distribution of this sum X , and find that if $V(r) \sim 1/r^{2+\epsilon}$, where $\epsilon > 0$, then $X(0)$ has a distribution law with a perfectly finite most probable value, while the distribution falls off as $1/X^3$ for large values of X . This latter fact shows, as can also be verified directly, that the mean value of X is infinite. Clearly this is merely the result of an infinitesimal fraction of atoms having a very large value. It can be shown that even these few large values are illusory, and that by suitably redefining the localized states in accordance with multiple-scattering theory⁷ the probability of such divergences is greatly reduced.⁸

In reference 2 the transition probability is calculated by taking the mean of (11) over all possible starting atoms. The resulting finite transition probability is therefore meaningless, as discussed above.

The case of $V(r) = A/r^3$, or normal dipolar spin diffusion, is a special one. In this case $X(0) \rightarrow \infty$ for all atoms, but the divergence is so extremely weak that $V_c(0)$ is finite. In fact, the distribution of $X(s)$ has the same form, but the most probable value is roughly

$$[X(s)]_{\text{M.P.}} \approx \frac{n^2 A^2}{W^2} \left[\sinh^{-1} \left(\frac{W}{2s} \right) \right]^2, \quad (12)$$

where n is the density of sites and W is the width of the

⁶ J. Holtmark, Ann. Physik 58, 577 (1919).

⁷ See, for instance, K. M. Watson, Phys. Rev. 105, 1388 (1957), also the references therein.

⁸ In this case this is particularly simple: near such an infinity $\text{Re}(V_c)$ will also be large so that the energy will simply be changed to a value at which $\text{Im}(V_c)$ is finite.

distribution of the E_j . This most probable value diverges as the square of $\ln s$ when $s \rightarrow 0$.

This case leads to a transport theory with decays slower than exponential. Since the divergence of (12) is caused by large values of r , single, rather long, jumps have an important effect on the transport process. The existence of this transport process for $V \sim 1/r^3$ provides a counter-example for those who may think the present theorem is self-evident for small enough V .

One may roughly estimate transport times by noting that, by the defining relationship (2), that value of s for which $s f(s)$ is of order unity is something like the inverse of a time of decay due to spin diffusion. Equation (12) shows that this time increases exponentially with W/nA for large values—for example, for Si donors at 10^{16} concentration we compute a rate of $\exp(-10^4) \text{sec}^{-1}$.

On the other hand, for $V \sim 1/r^{2+\epsilon}$ (or exponential, as it often is), the first term of perturbation theory leads to a vanishing rate of transport independent of V or W . Thus, if transport is to appear at all it must come in higher terms, and in fact it is easy to convince oneself that it can come only by a divergence of the whole series for V_c . Therefore it is of great importance to our theory to learn how to handle the sums of products

$$V_c(0) = \sum_{j,k,l,\dots,m \neq 0} V_{0j} \frac{1}{is - E_j} V_{jk} \frac{1}{is - E_k} \times V_{kl} \frac{1}{is - E_l} \dots V_{m0}, \quad (13)$$

which represent the possibility of successive virtual transitions until, at possibly some very great distance from site 0, a real, phase-destroying process can occur.

Our method for this problem, set forth in Sec. IV, involves both the idea of calculating a probability distribution rather than a mean for these terms, and also a modification of the multiple-scattering methods of Watson⁷ in order to eliminate certain troublesome repeated terms. We must do this elimination first.

Certain terms in (13) are apparently very large because of the fact that there is no prohibition on repeated indices. Suppose, for instance, that V_{jk} is particularly large and that both $is - E_j$ and $is - E_k$ are particularly small so that

$$\left| \frac{1}{is - E_j} V_{jk} \frac{1}{is - E_k} V_{kj} \right| > 1.$$

Then there is the possibility of a term like

$$V_{0l} \frac{1}{is - E_l} \dots V_{mj} \frac{1}{is - E_j} V_{jk} \frac{1}{is - E_k} V_{kj} \frac{1}{is - E_j} \dots \times V_{jk} \frac{1}{is - E_j} \dots \frac{1}{is - E_k} V_{kn} \dots \frac{1}{is - E_p} V_{p0}.$$

Such a term will get larger the more times we repeat V_{jk} . We can represent the terms of our series by closed diagrams through the various sites of the lattice, starting and ending at 0 (see Fig. 1), and this type of diagram involves a "ladder" which repeatedly runs back and forth from j to k . Physically, we can think of it as resulting from a pair of closely coupled atoms.

The technique of Watson^{7,9} shows us that we may eliminate all repeated indices in a self-consistent way by including in the energy denominator for atom k the perturbed energy $V_c(k)$ calculated from just such a series of terms as (13). A complicating factor is that $V_c(k)$ must be calculated from a series of diagrams which do not include *any* indices which have previously appeared before in the particular term of $V_c(0)$ we are calculating. That is, if we want the term

$$V_{03} \frac{1}{e_3} V_{32} \frac{1}{e_2} V_{21} \frac{1}{e_1} V_{10},$$

where for brevity we introduce the usual "propagator" notation

$$i_s - E_j - V_c(j) = e_j, \tag{14}$$

then the propagator e_2 , for instance, is given by

$$e_2 = i_s - E_2 - V_c^{0,1}(2) \\ = i_s - E_2 - \sum_{j,k,l \neq 0,1} \frac{1}{e_j} V_{2j} \frac{1}{e_l} V_{jl} \cdots V_{k2} \tag{15}$$

and again, each of the propagators in this series must be appropriately modified not to include either 0, 1, or any of the previous indices in the $V_c^{0,1}(2)$ series.

Thus we may now write

$$V_c(0) = \sum_{\substack{i \neq 0 \\ j \neq i, 0 \\ k \neq i, j, 0 \\ \dots \\ l \neq \dots, i, j, k, 0}} V_{0i} \frac{1}{i_s - E_k - V_c^{0,i}(k)} \\ \times V_{kj} \frac{1}{i_s - E_j - V_c^{0,i}(j)} V_{ji} \frac{1}{i_s - E_i - V_c^{0,i}(i)} V_{i0}. \tag{16}$$

All of this, of course, involves a self-consistent type of reasoning, since it is only if these series converge that we can find $V_c(j)$ in this way, and therefore that we can define the modified series. We say in defense that clearly we can always make the sum converge for large enough s , and also that the V_c 's in the higher terms, since they may have many forbidden indices, are more convergent than those we derive from them.

The prohibition of repeated indices has two useful consequences. The most obvious is to prevent extensive correlations between successive factors V/e of a given product. However, they also introduce a useful and

⁹ For this purpose, one could equally well use the method of E. Feenberg, Phys. Rev. 74, 206 (1948).

important correlation. Namely, suppose that one factor of our term, $V_{kl}/e_k^{(l \dots)}$, is particularly large. The V_c of the previous factor, say $V_{lm}/e_l^{(\dots)}$, will contain the term:

$$|V_{kl}|^2 / e_k^{(l \dots)}.$$

Therefore this previous factor would contain the large factor in its denominator, leading to a tendency to cancel. On a quantitative basis, first think of all V 's as having the same order of magnitude. Then, since the other terms of the denominator e_l will all be of order W or less, it is easy to see that we simply decrease the total unless

$$|V_{kl}|^2 / e_k \leq W,$$

or

$$|V_{kl}/e_k| \leq W/V. \tag{17}$$

W is the breadth of the distribution $P(E)$.

We shall use the limitation (17) in our later computations. We note that it is meaningless if V is small; but the work of Sec. IV will show that small values of V are never important. In any case the results do not depend sensitively on the existence of this limitation.

Actually, (17) is only the most important of an extensive system of correlations, since similar considerations hold for any group of factors starting from atom k and ending at atom l , if there is a distinct return path to atom k from l which has a finite factor. The fact that isolated large factors are not important,

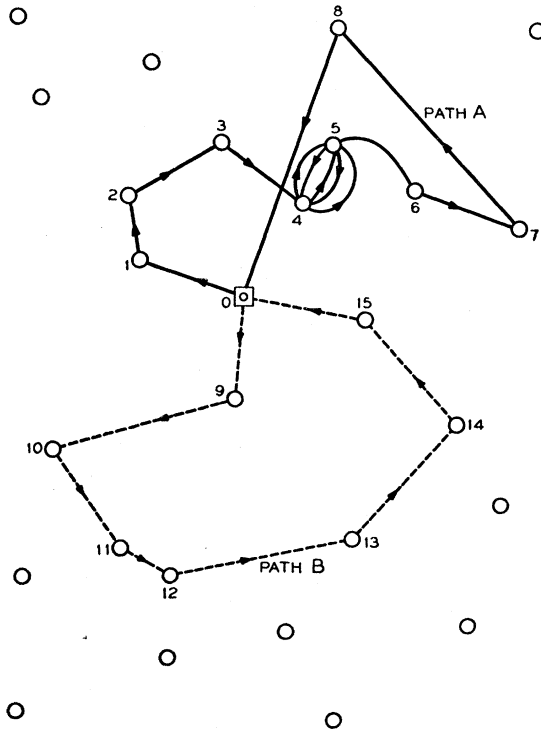


Fig. 1. Diagrams corresponding to terms in the perturbation expansion of V_c . (A) may be large and must be summed over; (B) is a legitimate term.

however, should apply even more strongly to groups of factors. We merely note that we will probably underestimate the limiting density, even using (17).

Our problem now is to study series of the form (16) in which the E 's, to a lesser extent the V_c 's, and possibly the V_{jk} 's are stochastic variables. We want primarily to find whether such series converge in some sense as $s \rightarrow 0$; if so, we have found a self-consistent set of localized wave functions. If the convergence limit on s is finite it may still be possible to calculate some transport properties, but that can be reserved for later work.

In the fourth section we discuss such series. The principle of this discussion is the following: since the terms T_L of the series having a given length (number of denominators, say) L , are themselves random variables, we try to find a distribution function for their values. This distribution can be expressed as a number distribution.

(Average number of terms of length L between T_L and $T_L + dT_L$) = $n(T_L)dT_L$. (18)

(It is necessary to use a number rather than a probability distribution because, when V extends to large r , the total number of terms of any length is infinite.)

The techniques involved in getting such a distribution are closely related to the Markoff method of random walk theory, although we use, for convenience, the Laplace transform and the convolution theory for it. We are able to get $n(T_L)$ explicitly in two cases. In both cases we make the unimportant simplification that the distribution function $P(E)$ of the energies E_j is flat:

$$P(E) = 1/W \quad \text{for} \quad -\frac{1}{2}W \leq E \leq \frac{1}{2}W, \quad (19)$$

$$P(E) = 0 \quad \text{for} \quad |E| > \frac{1}{2}W,$$

and we neglect the influence of V_c on the frequency denominators except for the limitation (17).

In the first case we assume that V_{jk} is finite only between "nearest neighbors," of which there are some finite number Z ; between these neighbors, it has a constant value V . Then the problem simplifies to finding the probability distribution of the product of denominators $P(\Pi_D)d\Pi_D$

$$\Pi_D = \frac{1}{e_1} \frac{1}{e_2} \frac{1}{e_3} \dots \frac{1}{e_L}. \quad (20)$$

Given $P(\Pi_D)$, we can use the idea of the "connectivity" from the percolation theory of Broadbent and Hammersley.¹⁰ The connectivity K for any given lattice with near-neighbor connections only is defined by exactly the relation we want:

(Number of nonrepeating paths of length L leading from any given atom) $\sim K^L$. (21)

¹⁰ S. R. Broadbent and J. M. Hammersley, Proc. Cambridge Phil. Soc. 53, 629 (1957). This work suggested some features of our approach to the present problem.

K is generally of order $Z-1$ to $Z-2$. Then obviously

$$n(T)dT = (K^L/V^L)P(T/V^L)dT. \quad (22)$$

A second case we are able to solve is that of the purely random lattice in which V falls off as some power of r . Unfortunately, we have to ignore the restriction of nonrepeating paths in this case, so that we rather badly overestimate the sum. On the other hand, this at least tells us whether or not large r 's are important since this restriction is not important for large jumps. Thus by this case we can show rigorously that $V \sim 1/r^{3+\epsilon}$ is the correct restriction on the range of V .

In each case we come out with an $n(T)$ of the following form:

$$n(T)dT = [F(K,W/V)]^L \frac{dT}{T^2} L(T), \quad (23)$$

where $L(T)$ is a slowly varying function relative to T . This form allows us to make use of the following result, which is implicit in (for instance) the theory of the Holtmark distribution. The probability distribution of the sum of a collection of random terms of random sign such as (23) is the same as the distribution of the single largest term (a) for values greater than or of the order of the most probable or median value of the sum; and (b) if $L(T)$ is increasing, or decreasing no more rapidly than $T^{-1/2}$. This is essentially the same as the theorem that the force on a dipole in an unpolarized gas, or on an electron in a discharge, comes primarily from the nearest neighbor.

Since $L(T)$ obeys this condition very well in both cases, at least for reasonable values of V/W , etc., we may immediately get the critical values of the parameters from (23). We know now that, if

$$\Sigma = \sum_{n=1}^{K^L} (\pm T_n) \quad (24)$$

(where we use for clarity the case of a finite number of neighbors), then

$$P(\Sigma)d\Sigma \sim F^L(K,W/V) \frac{d\Sigma}{\Sigma} L(\Sigma). \quad (25)$$

First we find $(W/V)_0$ to satisfy

$$F^L(K, (W/V)_0) L(1) = 1. \quad (26)$$

If (W/V) has a value even very slightly greater than this, the most probable value of Σ will be small of order e^{-L} and the probability of a value = 1 will also be of order e^{-L} . Now we consider $L \rightarrow \infty$: The number of Σ 's only increases as L , while with probability $\sim 1 - e^{-L}$ their value is less than e^{-L} . Therefore the series converges almost always if

$$W/V > (W/V)_0, \quad (27)$$

which is the desired criterion. In Sec. IV the critical values will be discussed numerically. A typical estimate

would be $(W/V)_0=26$ for $K=4.5$ (about correct for the simple cubic lattice).

With this result the theorem is established.

III. PROBABILITY DISTRIBUTION OF THE FIRST TERM OF V_s

Before Eq. (11) we related the transition rate to a certain quantity $X(s)$, and showed that if $X(s)$ remains finite as $s \rightarrow 0$ no real transport takes place, in the first order of perturbation theory. Now $X(s)$ is a sum over all possible single jumps:

$$X(s) = \sum_k \frac{|V_{0k}|^2}{s^2 + E_k^2}. \quad (28)$$

The probability distribution of such a sum is best calculated by the Markoff method,¹¹ as modified by Holtmark.⁶ In this method we find the Fourier transform of the probability distribution $P(X)$:

$$P(X) = \int_{-\infty}^{\infty} e^{ixX} \varphi(x) dx, \quad (29)$$

where

$$\varphi(x) = \exp \left\{ -n \left\langle \int \left[1 - \exp \left(ix \frac{V^2(r)}{E^2 + s^2} \right) \right] d\tau \right\rangle \right\}. \quad (30)$$

The average is to be taken over the probability distribution of E , $P(E)$, and n is the density of sites.

Let us write out the important integral in the exponent of (30):

$$I = \int_{-\infty}^{\infty} P(E) dE 4\pi \int_0^{\infty} r^2 dr \left[1 - \exp \left(\frac{ixV^2(r)}{E^2 + s^2} \right) \right]. \quad (31)$$

The behavior of $P(X)$ for large X depends on the behavior of I for sufficiently small x . Let us first consider the case $s=0$. Now for small enough x , and a finite E (say of order W), the exponential $\exp(ixV^2/E^2)$ can be expanded in a power series in x , and the integration over r done (so long as V is finite and falls off faster than r^{-3}) to obtain terms which go as x^1 or higher powers for small x . Thus only the behavior for small E is important, and we can neglect the variation of $P(E)$, replacing it by a constant, $1/W$. Then

$$\begin{aligned} I &\cong \frac{4\pi}{W} \int_0^{\infty} r^2 dr \int_{-\infty}^{\infty} dE \left\{ 1 - \exp \left[\frac{ixV^2(r)}{E^2} \right] \right\} \\ &= \frac{4\pi}{W} (x)^{\frac{1}{2}} \int_0^{\infty} r^2 dr V(r) \int_{-\infty}^{\infty} du \left[1 - \exp(i/u^2) \right] \\ &= 2 \left(\frac{x}{i} \right)^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right) \frac{\langle V(r) \rangle}{W}. \end{aligned} \quad (32)$$

¹¹ S. Chandrasekhar, *Revs. Modern Phys.* **15**, 1 (1943). I am indebted to L. R. Walker for suggesting the use of the Markoff method here.

The integration here depends on the more stringent condition $V \sim 1/r^{3+\epsilon}$ for large r . The probability distribution, which will be valid for large X at least, is familiar from line-broadening theory¹²:

$$P(X) = \frac{n\langle V \rangle}{WX^{\frac{1}{2}}} \exp \left[- \left(2n\Gamma\left(\frac{1}{2}\right) \frac{\langle V \rangle}{W} \right)^2 \frac{1}{X} \right]. \quad (33)$$

For large X , this falls off as $X^{-\frac{1}{2}}$, as stated in Sec. II: while the mean of X is divergent, the probability that X is larger than some value X_0 decreases as $X_0^{-\frac{1}{2}}$. Thus, for any given starting atom n , the renormalization constant K may be large, but the probability that τ is finite is exactly zero.

We see that none of these conclusions are valid for the case $V \propto r^{-3}$ for large r . In this case a finite s must be retained. Again looking at the integral (31), let us substitute $V(r) = A/r^3$ for all r .¹³ Then we do the integration over r first:

$$\begin{aligned} I &= \frac{4\pi}{3} \int_{-\infty}^{\infty} P(E) dE \int_0^{\infty} d(r^3) \left\{ 1 - \exp \left[\frac{ixA^2}{r^6(E^2 + s^2)} \right] \right\} \\ &= \frac{4\pi A}{3} \left(\frac{x}{i} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} \frac{P(E) dE}{(E^2 + s^2)^{\frac{3}{2}}}. \end{aligned} \quad (34)$$

We see that I indeed has a logarithmic singularity as $s \rightarrow 0$. A simple case for which we can evaluate the integral (34) is the flat distribution (19) of width W for which

$$\frac{W}{2} I = \frac{4\pi A}{3} \left[\sinh^{-1} \left(\frac{W}{2s} \right) \right] \left(\frac{x}{i} \right)^{\frac{1}{2}}. \quad (35)$$

This leads to the probability distribution of the sum X :

$$\begin{aligned} P(X) &= \left(\frac{4\sqrt{2}nA}{3\pi^{\frac{1}{2}}W^{\frac{1}{2}}} \right) \frac{1}{X^{\frac{1}{2}}} \sinh^{-1} \left(\frac{W}{2s} \right) \\ &\times \exp \left\{ - \left[\frac{8\pi nA}{3W} \sinh^{-1} \left(\frac{W}{2s} \right) \right]^2 \left(\frac{1}{X} \right) \right\}, \end{aligned} \quad (36)$$

which was discussed in Sec. II.

IV. DISTRIBUTIONS OF HIGH-ORDER TERMS IN THE PERTURBATION THEORY

In order to simplify the later manipulations, and to make closer contact with the Watson theory, it will be useful to expand some of the formalism of the problem.

Equation (4) presents our basic equations of motion:

$$f_j(s) = \frac{i\delta_{0j}}{is - E_j} + \sum_k \frac{1}{is - E_j} V_{jk} f_k(s).$$

¹² H. Margenau, *Phys. Rev.* **48**, 755 (1935).

¹³ We make this substitution since there exists some r_0 beyond which $V - (A/r^3) \simeq 0$. We point out that $\int_0^{r_0}$ leads to terms in I of identical form with those we have already found, and thus independent of s .

Let us consider simultaneously all possible initial starting atoms 0. The $f(s)$'s which result will form a matrix f_{lm} , of which our f 's are the particular row f_{j0} . It is simple to introduce as the "Green's function" the matrix

$$(j|W|k) = -i f_{jk}(s), \tag{37}$$

which satisfies the equation

$$(j|W|k) = \frac{\delta_{jk}}{is - E_j} + \sum_l \frac{1}{is - E_j} V_{jl} (l|W|k); \tag{38}$$

or, if one introduces the matrix

$$(j|a|k) = \delta_{jk}(is - E_j),$$

then W satisfies

$$W = -\frac{1}{a} + \frac{1}{a} VW. \tag{39}$$

The Møller wave matrix Ω is defined also, as

$$\Omega = Wa, \tag{40}$$

and satisfies the equation

$$\Omega = 1 + \frac{1}{a} V\Omega. \tag{41}$$

The general philosophy of the multiple-scattering theory is to try to separate and identify, in these matrices Ω or W , effects which are "coherent" in the sense that they involve perturbations in which, finally, the system returns coherently to the initial state, from effects which involve real (as opposed to virtual) transitions to other states. Another way to put it is that this is an attempt to define a new set of states, once the perturbation V is applied, which correspond in some sense to the unperturbed states, and thus to have left over only the effects of whatever real transitions may occur. In principle this is exactly what we are attempting in this paper.

The coherent effects are mainly included in a diagonal coherent wave matrix Ω_c , and the rest of the problem is contained in the "model operator" F :

$$\Omega = F\Omega_c, \tag{42}$$

where we want F to satisfy an equation like

$$F = 1 + \frac{1}{a - V_c} P V F, \tag{43}$$

where P is an operator which prevents in some way the repetition of indices in the perturbation series for F , while V_c is a correction which must therefore be made to the energy a .

It is perhaps simplest just to go ahead and show what must be done. The solution of Eq. (41) may be

written out as a perturbation series by direct iteration:

$$(j|\Omega|k) = \delta_{jk} + \frac{1}{a_j} V_{jk} + \sum_l \frac{1}{a_j} \frac{1}{a_l} V_{jl} V_{lk} + \sum_{l,m} \frac{1}{a_j} \frac{1}{a_l} \frac{1}{a_m} V_{jl} V_{lm} V_{mk} + \dots \tag{44}$$

A very direct way to eliminate repeated indices is as follows: take any given term of (44) and, starting from the right, look through until we find the first repetition of the index k . Between these two repetitions comes a factor:

$$\left(\frac{1}{a_k} \frac{1}{a_l} \dots V_{mk} \right).$$

Next we continue to look from right to left and find another similar factor, etc., until we find all such factors and no more k indices occur. Now everything which remains to the left will also appear multiplied by the factors we have found in all other possible combinations and repetitions, and in fact by all possible such factors in all such combinations. Summing all such factors, we get for that series which comes to the right of the last repetition of index k :

$$\begin{aligned} & \left[1 + \frac{1}{a_k} V_{kk} + \left(\frac{1}{a_k} V_{kk} \right)^2 + \dots + \sum_l \frac{1}{a_{kl}} V_{kl} \frac{1}{a_l} V_{lk} \right. \\ & \quad \left. + \left(\sum_l \frac{1}{a_k} \frac{1}{a_l} V_{kl} V_{lk} \right)^2 + \dots \right] \\ & = \left[1 - \left(\frac{1}{a_k} V_{kk} + \sum_l \frac{1}{a_k} \frac{1}{a_l} V_{kl} V_{lk} + \dots \right) \right]^{-1} \\ & = \frac{a_k}{a_k - V_c(k)} = \frac{a_k}{e_k} = (\Omega_c)_k, \tag{45} \end{aligned}$$

where $V_c(k)$ is defined as in (13). The corresponding term of Ω is then in the form

$$\frac{1}{a_j} \frac{1}{a_l} V_{jl} V_{lm} \dots \frac{1}{a_n} V_{nk} (\Omega_c)_k.$$

We now begin the same process as in (45), looking for repetitions of the index n which appear next. We collect together all such terms, and finally find that we can replace a_n by $a_n - V_c^k(n)$, where

$$\begin{aligned} V_c^k(n) = & V_{nn} + \sum_{l \neq k, n} \frac{1}{a_l} V_{nl} V_{ln} \\ & + \sum_{l, m \neq k, n} \frac{1}{a_l} V_{nl} V_{lm} \frac{1}{a_m} V_{mn} \dots \tag{46} \end{aligned}$$

This process may be repeated until we come to the end

of the given term. Thus we have successfully expanded Ω as

$$\begin{aligned}\Omega &= F\Omega_e, \\ \Omega_e &= a/e, \\ F_{j\neq k} &= -\frac{1}{e_j}V_{jk} + \sum_{l\neq k, i} \frac{1}{e_l}V_{jl}V_{lk} \\ &\quad + \frac{1}{e_j} \sum_{l\neq m, k, i} V_{jl} \sum_{e_l} \frac{1}{e_l} \sum_{m\neq j, k} V_{jm}V_{mk} + \dots\end{aligned}\quad (47)$$

The final step is obvious: in each of the V_e 's in (45) itself we eliminate repeated indices in an exactly similar manner, obtaining expressions like (16) for V_e .¹⁴

Now the usefulness of such expressions in the usual multiple-scattering theory comes only from the fact that the limitations on the sums are unimportant, since there are an infinite number of V_{jk} 's starting from any j , each being small. In our case we have a different kind of fortunate circumstance which allows us to get around this problem; namely, all of the quantities of the theory are stochastic variables, so that all we really wish to know is the distribution function of the e_j 's, which except for the restriction (17) is practically that of the E_j 's unless V_e is quite large. Even if V_e were large, one might study it as a stochastic variable.

We now have the problem of calculating probability distributions for products such as the terms of (47). Because the only question is that of convergence we are interested only in the terms of very high order, that is, those of order L with $L \gg 1$. We call such a term T :

$$\begin{aligned}T_{jk}^L &= \frac{1}{e_j}V_{jl} \frac{1}{e_l}V_{lm} \dots V_{nk} \\ &= \exp\{\ln V_{jl} + \ln V_{lm} + \dots (L \text{ terms})\} \\ &\quad - [\ln e_j + \ln e_l + \dots (L \text{ terms})].\end{aligned}\quad (48)$$

In Sec. II we presented the two cases in which it has been possible to calculate explicitly the number distribution of T . We take up the simplest first: the case in which V is a constant, so that the only stochastic element in (48) is the denominators. In this case

$$T = V^L \prod_{i=1}^L \frac{1}{e_i}.\quad (49)$$

To find the convergence limits, we go to $s=0$ immediately. Then clearly T has random sign, which must be taken into account later in summing the T 's.

As we discussed before, V_e is important only in that it causes a certain restriction (17) on the magnitude of the separate factors of the product; otherwise we shall

¹⁴ I am indebted to P. A. Wolff for many helpful discussions on the above, and in particular for pointing out that (43) is true only in the sense of the perturbation series.

make the approximation of neglecting it altogether as an unimportant correction to the stochastic variable E_j . Since we are for simplicity confining ourselves to the flat distribution $P(E)=1/W$ of Eq. (19), we can express (17) semiquantitatively. The smallest denominator e_l which will not seriously affect the probability distribution of subsequent factors we define to be $\Delta/2$. The quantity $\Delta/2$ satisfies the criterion that the contribution to V_e from it must be less than the maximum possible E_j :

$$V^2/(\frac{1}{2}\Delta) < \frac{1}{2}W,$$

or

$$\Delta > 4V^2/W.\quad (50)$$

Our approximation is simply to take this as a lower limit on e and modify (19) to

$$P(e_j) = \frac{1}{W-\Delta}, \quad \frac{1}{2}\Delta < |e_j| < \frac{1}{2}W;\quad (51)$$

$$P(e_j) = 0, \quad |e_j| < \frac{1}{2}\Delta \quad \text{or} \quad |e_j| > \frac{1}{2}W.$$

We shall find that the use of (51) changes our convergence limits by less than a factor of 2, in spite of its apparent importance in eliminating singular factors; this is our justification for the crude approximation.

We now take up the question of the probability distribution of T . Let us define the variable S as

$$\left| \prod_{j=1}^L \frac{1}{e_j} \right| = e^{S/(\frac{1}{2}W)^L}.\quad (52)$$

The range of S starts from zero, so that we can apply a Laplace transformation to its probability distribution:

$$\begin{aligned}F_L(p) &= \int_0^\infty e^{-pS} P(S) dS \\ &= \left\langle \prod_{j=1}^L \exp\{p[\ln e_j - \ln(\frac{1}{2}W)]\} \right\rangle.\end{aligned}\quad (53)$$

Since all the e_j 's are independent variables, the average in (53) is just the product of the separate averages. Thus

$$F_L(p) = \left[\frac{1 - (\Delta/W)^{p+1}}{p+1} \frac{W}{W-\Delta} \right]^L \quad (p > -1).\quad (54)$$

For regions of S in which we may neglect Δ , this is very easily inverted:

$$\Delta \cong 0: \quad F_L(p) \cong 1/(p+1)^L, \quad P(S) = e^{-S} S^{L-1} / \Gamma(L).\quad (55)$$

Since

$$T = V^L / (\frac{1}{2}W)^L e^{S/(\frac{1}{2}W)^L},$$

to our order of approximation we have

$$P(T) dT \cong \left(\frac{2eV}{W} \right)^L \left[\frac{\ln T}{L} - \ln \left(\frac{2V}{W} \right) \right]^L \frac{dT}{T^2},\quad (56)$$

which shows that it is of the form (23).

It is interesting to note that while (55) is a reasonably narrow distribution, satisfactorily obeying the central limit theorem, the fact that the quantity of interest is exponential in S transfers our attention to what, in (55), appears to be the extreme tail of the distribution. This is a characteristic of this problem. On the other hand, our task is simplified, in that nothing smaller than factors exponential in L affects our results.

Without neglecting Δ , we can get the full distribution with sufficient accuracy from (54) by applying the inversion formula for the Laplace transform,

$$P(S) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} F_L(p) e^{pS} dp, \quad (57)$$

and using the method of steepest descents. Writing out (57) in appropriate form for this method, we obtain

$$P(S) = \frac{1}{2\pi} \int_{-i\infty}^{i\infty} \exp \left\{ pS - L \left[\ln(p+1) + \ln \left(1 - \frac{\Delta}{W} \right) - \ln \left[1 - \left(\frac{\Delta}{W} \right)^{1+p} \right] \right] \right\}. \quad (58)$$

First notice that as p becomes large and positive, the exponent will approach $+\infty$ unless $S < 0$. If $S < 0$, we can find a path via $p \rightarrow +\infty$ and $P(S) = 0$, as it should be. Similarly, as $p \rightarrow -\infty$, the S term dominates if $S > L \ln(W/\Delta)$, so that, correctly, $P(S > L \ln(W/\Delta)) = 0$. Within these limits, there is a saddle point for finite p which may be found by differentiating the exponent.

At the point $1+p=0$, the exponent changes character, from not depending on Δ/W above this point to depending primarily on it below. It is instructive to expand the exponent about this point:

$$-L \left\{ \ln(1+p) + \ln \left(1 - \frac{\Delta}{W} \right) - \ln \left[1 - \left(\frac{\Delta}{W} \right)^{1+p} \right] \right\} \cong -L \left\{ \ln \left(1 - \frac{\Delta}{W} \right) - \ln \left[\ln \frac{W}{\Delta} - \frac{1+p}{2} \left(\frac{W}{\Delta} \right)^2 \dots \right] \right\}. \quad (59)$$

As we see, this point is actually not a singularity. Taking derivatives, we find that the condition that the saddle be here is:

$$S|_{p=0} = -L = \frac{1}{2} L \ln(W/\Delta),$$

and that at this point

$$P \left(\frac{1}{2} L \ln \frac{W}{\Delta} \right) \cong e^{-S} \left(\frac{\ln(W/\Delta)}{1-\Delta/W} \right)^L. \quad (60)$$

As S gets appreciably larger than this value, $F_L(p)$ and the exponent again become simple when we can

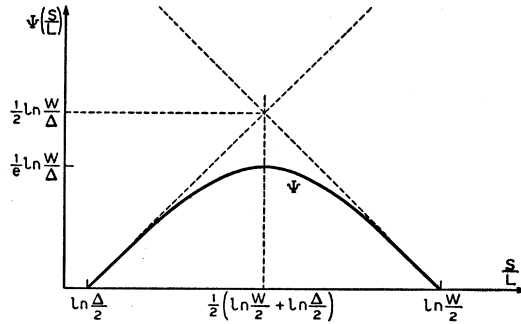


FIG. 2. The function ψ of Eq. (64), giving the slowly varying part of $P(S)$.

neglect 1 relative to $(\Delta/W)^{p+1}$. Here it becomes necessary to continue the logarithms in (58) to negative values of their arguments, but this can be done by referring to (54) and noticing that we must change the sign of $1+p$ and $-(\Delta/W)^{p+1}$ simultaneously. Then the saddle-point condition is

$$S - \frac{L}{1+p_0} + L \ln \frac{\Delta}{W} = 0, \quad (61)$$

and the probability of such values of S is

$$P(S) \cong e^{L-S} \left(\ln \frac{W}{\Delta} - \frac{S}{L} \right)^L / \left(1 - \frac{\Delta}{W} \right)^L. \quad (62)$$

The results (56), (60), and (62) may be summarized in the following way:

$$P(S) = \left(\frac{e}{1-\Delta/W} \right)^L e^{-S} [\psi(S/L)]^L, \quad (63)$$

where ψ is a slowly varying function which may be estimated in each of the three regions:

- I: $\frac{1}{2} L \ln(W/\Delta) \gg S > 0, \quad \psi(S/L) \cong S/L;$
- II: $S \cong \frac{1}{2} L \ln(W/\Delta), \quad \psi(S/L) \cong \ln(W/\Delta)/e;$
- III: $L \ln(W/\Delta) > S \gg \frac{1}{2} L \ln(W/\Delta), \quad \psi(S/L) \cong \ln(W/\Delta) - (S/L).$

The function ψ is easily plotted approximately from these results and is shown in Fig. 2. For the probability distribution of T , we obtain

$$P(T) dT = \frac{dT}{T^2} \left(\frac{2eV}{W} \right)^L \times \psi^L \left[\frac{\ln T}{L} - \ln \frac{2V}{W} \right] / \left(1 - \frac{\Delta}{W} \right)^L. \quad (65)$$

Except for the most remote parts of region III, (65) is of the form (23) and satisfies our condition.

We have not explored very carefully the corrections to the saddle-point method, both because we need only

terms of the order e^L , and because the results are clearly in accordance with expectations.

Before studying (65) numerically, let us go on to the second case which can be solved; namely,

$$V = V_0/r^{3N}. \quad (66)$$

We find a solution only in the limited sense that various crude approximations are made for small r ; what we try to do is to assure ourselves that the region of large r and small V is not important, in spite of the extra mathematical difficulties to which it leads.

These "crude approximations" are threefold: (a) We ignore the fact that a path leading to an atom from its nearest neighbor must leave it via a further neighbor—i.e., in each factor we allow V to be randomly distributed, ignoring the favorable correlations caused by the restriction to nonrepeating paths. (b) We ignore (17) and thus can use (55) and (56) as the distribution function of the denominators, again overestimating the effects of large V 's without disturbing the small V region. (c) We limit $|V|$ simply by introducing a minimum radius a ; the maximum V is then

$$V < V_m = V_0/a^{3N}. \quad (67)$$

Now we study the distribution of terms of the form

$$T = \prod_{j=1}^L V_j / \prod_{j=1}^L e_j. \quad (68)$$

The distribution of the V 's may be approximated by using a perfectly random distribution of neighbor distances:

$$n(r)dr = 4\pi\rho r^2 dr,$$

where ρ is the density of sites. From this and (67) the distribution of V is immediately

$$\begin{aligned} n(V)dV &= \frac{4\pi\rho}{3N} V_0^{1/N} \frac{dV}{V^{1+1/N}}, \quad V < V_m \\ &= 0, \quad V > V_m. \end{aligned} \quad (69)$$

The numerator in (68) may take on values from V_m^L to zero; the inverse of the denominator, from ∞ to a certain minimum. Again the mathematics is more familiar if we study the logarithm, which is a variable extending from $-\infty$ to ∞ . Thus it is necessary to use now the bilateral Laplace transform. This causes no difficulty in (53), (55), and (56) since $S > 0$ anyhow; we simply reinterpret these as the corresponding bilateral formulas.

The logarithmic variable Σ , corresponding to S , we define by

$$\prod_{j=1}^L V_j = V_m^L e^{\Sigma}, \quad (70)$$

and the quantities (68) whose distribution we want are

$$T = \frac{V_m^L}{(W/2)^L} e^{S+\Sigma}. \quad (71)$$

Thus, once we have the distribution of Σ , we can find the distribution of $S+\Sigma$ by convolution of the bilateral Laplace transforms, and thence find that of T .

The required Laplace transform is

$$\begin{aligned} \varphi_L(p) &= \int_{-\infty}^{\infty} P(\Sigma) e^{-p\Sigma} d\Sigma \\ &= \left\{ \frac{4\pi\rho a^3}{3N} \int_0^1 \frac{(V/V_m)^{-p} d(V/V_m)}{(V/V_m)^{1+1/N}} \right\}^L, \end{aligned} \quad (72)$$

$$\varphi_L(p) = \left(\frac{a^3}{Nr_s^3} \right)^L \left\{ -\frac{1}{p+1/N} \right\}^L, \quad p < -1/N. \quad (73)$$

Here we have defined

$$\frac{4}{3}\pi r_s^3 = 1/\rho. \quad (74)$$

The transform of the denominators, $F_L(p)$, was convergent for $p > -1$ [Eq. (55)]. Thus for $N > 1$, or forces falling off more rapidly than $1/r^3$, the two transforms have a common strip of convergence. This is the criterion on range we already expected from Sec. III. Now let

$$X = S + \Sigma,$$

$$\int_{-\infty}^{\infty} n(X) dX e^{-pX} = \psi_L(p). \quad (75)$$

Also, by the convolution theorem,

$$\begin{aligned} \psi_L(p) &= \varphi_L(p) F_L(p), \quad -1 < p < -1/N \\ &= \left(\frac{a^3}{Nr_s^3} \right)^L \left\{ -\frac{1}{(1+p)(p+1/N)} \right\}^L. \end{aligned} \quad (76)$$

The inversion of this is simple if we apply the shifting operator to p , bringing the origin to the center of the convergence strip:

$$p' = p + \frac{1}{2}(1+1/N),$$

$$\psi_L(p') = \int_{-\infty}^{\infty} \exp(-p'X) n'(X) dX, \quad (77)$$

$$n'(X) = n(X) e^{\frac{1}{2}X(1+1/N)}.$$

Then

$$\psi_L(p') = \left[\frac{a^3}{Nr_s^3} \frac{1}{\frac{1}{4}(1-1/N)^2 - p'^2} \right]^L.$$

The standard inversion formula now gives us

$$n'(X) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{dp' \exp(p'X)}{[\frac{1}{4}(1-1/N)^2 - p'^2]^L} \left[\frac{a^3}{Nr_s^3} \right]^L. \quad (78)$$

This integral can be expressed in terms of the Bessel function $K_{L-\frac{1}{2}}$ (which is actually a polynomial in elementary functions) by Basset's formula,¹⁵ which in

¹⁵ G. N. Watson, *Bessel Functions* (Cambridge University Press, Cambridge, 1944), p. 172.

turn can be expanded by the well-known asymptotic expansion for large order.¹⁶ The result, for our purposes, could be foreseen by realizing that in (78) practically all the contribution will come from $p'=0$. Using (77), and neglecting unimportant constant factors, it is

$$n(X) = \left[\frac{4a^3}{Nr_s^3(1-1/N)^2} \right]^L \exp[-\frac{1}{2}(1+1/N)X]. \quad (79)$$

Now we get the number distribution of the terms T , using (71):

$$n(T)dT = \frac{dT}{T^{\frac{1}{2}+1/2N}} \left[\frac{4a^3}{Nr_s^3(1-1/N)^2} \right]^L \left[\frac{2V_m}{W} \right]^{\frac{1}{2}L(1+1/N)}. \quad (80)$$

Again we find a distribution of the form (23), satisfying very well the condition that the distribution of large values be that of the largest term.

Now the only remaining task to complete the discussion is to study the criterion for localization numerically, using (23) and (26). First we shall deal very briefly with the unrealistic second case. Upon using (80), Eq. (26) becomes

$$1 = \left[\frac{4a^3}{Nr_s^3(1-1/N)^2} \right]^L \left[\frac{2V_m}{W} \right]^{\frac{1}{2}L(1+1/N)}.$$

It is interesting to put this in the following form:

$$\frac{V_0}{\left\{ r_s^2 a^{N-1} \left[\frac{2}{N^{\frac{1}{2}} - 1/N^{\frac{1}{2}}} \right]^{4/3} \right\}^{3N/(N+1)}} = \frac{W}{2}. \quad (81)$$

The $3N$ th root of the denominator on the left can be thought of as an effective radius of interaction; namely, contributions from much greater distances are certainly of no importance. Except when N is very close to unity this radius is smaller than the mean nearest-neighbor distance ($\approx 0.8r_s$) and strongly dependent upon a . This tells us two things: first, that the infinite range of the potential is, unless $N \approx 1$, of no importance, and the important interactions are with close neighbors; and second, that this particular calculation, which did not take into account the important correlations of near neighbors introduced by the restriction to nonrepeating paths, is of no direct value.

An order-of-magnitude guess at the correct answer to this problem might be gotten by inserting something like r_s for a in (81); this eliminates the false effect of single, very close neighbors, which by the restriction to nonrepeating paths should make no contribution. We see then that $r_{\text{eff}} \approx r_s$, or approximately the mean nearest-neighbor distance. The resulting V/W is rather surprisingly large and tends to explain Feher's result that even when a considerable fraction of the atoms

have a close neighbor for which $V > W$, there is still little or no spin diffusion.

These difficulties with correlations confine our quantitative work to the one case in which this problem is solved for us, at least in principle, by Hammersley's work: The regular lattice with a finite number Z of equal interactions.

Here we apply very directly the discussion of Eqs. (22)-(26): that is, we find the probability distribution of the sum of K^L terms from that of the largest term, and thence find a critical $(W/V)_0$ below which all the higher terms are exponentially small with exponentially large probability. The criterion for applicability of the basic theorem in the case of the present distribution (65) turns out to be

$$\ln \frac{2V}{\Delta} - \frac{\ln T}{L} < 1, \quad (82)$$

which we can show to be true at the critical V/W in all cases.

Equation (26) for the critical W/V is, when one uses (65) and takes the L th root,

$$1 = \frac{2eK}{1 - (\Delta/W)_e} (V/W)_e \psi \{ -\ln [2(V/W)_e] \}. \quad (83)$$

Since we are really quite uncertain as to exactly how to take the correlations into account, we shall solve (83) first in the case of no correlation, which gives us an upper limit on $(W/V)_0$, as well as in the case of a finite Δ given by (50).

I. *Upper limit.*—Here we assume that $\Delta = 0$ and that ψ for region I is always correct, so that

$$eK \ln(W/2V)_{\text{u.l.}} = (W/2V)_{\text{u.l.}}. \quad (84)$$

A rough solution by iteration is clearly $(W/2V)_{\text{u.l.}} = eK \ln(eK)$. More accurate solutions can be obtained by plotting (84) and are given in Fig. 3. The values are rather large; for instance, at $K = 4.5$, approximately correct for the simple cubic lattice, $(W/2V)_{\text{u.l.}} = 45$. The rather large effect of increasing connectivity is interesting.

II. *Lower limit.*—Next let us do the calculation taking into account the approximate lower limit (50). It turns out that in this case the T of interest is in region II. In fact, with the definition (50),

$$-\ln(2V/W)_e = \frac{1}{2} \ln(W/\Delta),$$

so that S is exactly correct for region II. This leads to

$$\frac{2K \ln(W/2V)_e}{1 - (4V^2/W^2)_e} = (W/2V)_e. \quad (85)$$

Except for the correction in the denominator, which is negligible for all but the smallest values of K , this is the same equation as before with 2 replacing e . Its

¹⁶ J. W. Nicholson, Phil. Mag. 20, 938 (1910).

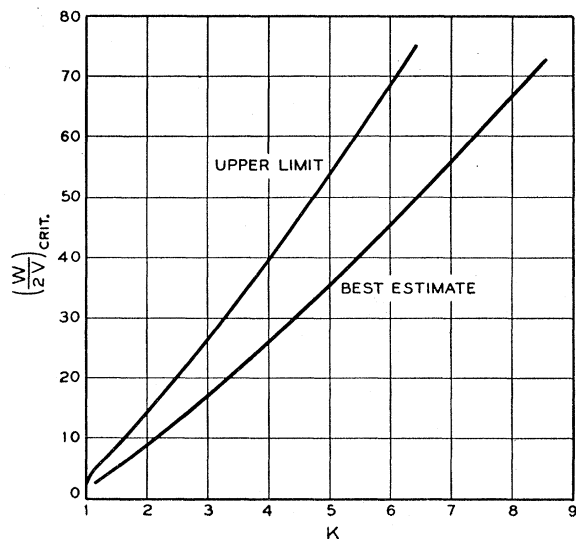


FIG. 3. Numerical estimates for the critical $W/2V$, the ratio of line width to interaction, for transport, plotted against connectivity K . The upper curve is a quasi-exact upper limit; the lower one is our best estimate.

solution is the lower curve of Fig. 3. It is very unlikely that (85) is accurate for $K \sim 1$, so no plot has been made in this region. This concludes our estimates of the critical ratio for transport.

V. MEANING OF CONVERGENCE OF THE PERTURBATION THEORY

The results of Secs. III and IV may be stated simply as follows:

(1) The first few terms of perturbation theory are convergent in the sense that $\text{Im}(V_e)$ is finite with probability unity, at any particular randomly chosen point on the energy axis.

(2) The terms of order L or higher are, at any particular random point on the energy axis, smaller than $e^{-\epsilon L}$ with probability $1 - e^{-\epsilon L}$, under the appropriate conditions.

We discuss here the question of what this tells us about the actual perturbed energy states. Our conclusion is that we can show that a typical perturbed state is localized with unity probability; but that we cannot prove that it is possible to assign localized perturbed states a one-to-one correspondence with localized unperturbed states in any obvious way, so that perhaps with very small probability a few states may not be localized in any clear sense.¹⁷

Equation (37) introduces the "Green's function" $\langle j|W(s)|k\rangle = -if_{jk}(s)$. Let us define

$$s = \sigma + iE. \quad (86)$$

According to the meaning of $f_{jk}(s)$, we can find the

¹⁷ Our criterion for nontransport is the same as that of L. Van Hove [Physica 23, 441 (1957); see especially Sec. 4] translated into probability language. Van Hove's resolvent R_t is our Green's function and his G_t our V_e .

final amplitudes on the various atoms by the prescription:

$$\lim_{\sigma \rightarrow 0} i\sigma \langle j|W|k\rangle = \bar{a}_{jk}(E), \quad (87)$$

where $\bar{a}_{jk}(E)$ is the amplitude at energy E and on atom j of the wave function which initially was unity on atom k . The total amplitude on atom j is the sum of squares of (87) over all energies E which are exact eigenenergies of the problem. Finally, W obeys the usual definition of the Green's function:

$$W = \frac{1}{iS - H_0 - V} = \frac{1}{i\sigma + E - H}. \quad (88)$$

[This is just (39).]

The scheme of the multiple-scattering method is to replace V in (88) by a number V_e , which can, of course, only be done in the diagonal elements:

$$\langle j|W|j\rangle = \frac{1}{i\sigma + E - E_j - V_e^j(i\sigma + E)}. \quad (89)$$

Then the general elements of W are obtained by using the model operator F of (42) and (47):

$$\langle j|W|k\rangle = \langle j|F|k\rangle \langle k|W|k\rangle. \quad (90)$$

It is important that in the expansion (47) of F the particular denominator occurring in (89) never appears. Let us start by thinking of a large but finite system, so that we know the energy levels are not a continuum. Then the exact energies of the problem are clearly the poles of W , which occur at

$$E - E_j - V_e^{(j)}(E) = 0. \quad (91)$$

This is just the expression one gets in Brillouin-Wigner perturbation theory. At such an energy, the amplitude at $t \rightarrow \infty$ is given by

$$\bar{a}_{jj}(E) = \lim_{\sigma \rightarrow 0} \frac{i\sigma}{i\sigma - i \text{Im}[V_e^{(j)}(i\sigma + E)]}, \quad (92)$$

so that if $\text{Im} V_e$ converges and this convergence remains as $N \rightarrow \infty$, this is finite, not zero as in usual transport theory.

Thus our proof that, at an arbitrarily chosen point E , $\text{Im} V_e$, or any of the quantities of the theory, converge, and (as we could also show) $\langle j|F|k\rangle$ falls off rapidly with distance, would be complete if (91) furnished us with only one solution per E_j , and if this solution were displaced by a finite and indeterminate amount from the poles of V_e or F . Unfortunately, this is only true of some of the solutions of (91), those which correspond to nearby localized functions; but because (91) is a Brillouin-Wigner result it also contains all the other eigenstates as well. We have to show, of those distant from atom j , that their energies are *not* random relative to the poles of $(V_e^{(j)})$, but instead are specially related to V_e so that \bar{a}_{jj} and \bar{a}_{jk} are decreasing functions of distance.

Why this is so is just as apparent from the simplest second-order theory as from the whole sum. Let us write (91) to second order:

$$E - E_j - \sum_k \frac{|V_{jk}|^2}{E - E_k} = 0.$$

Clearly this has a solution $E_0(j)$ near E_j at which the sum takes on the value

$$\sum_k \frac{|V_{jk}|^2}{E_j - E_k},$$

approximately. There is no close relation between $E_0(j)$ and any of the poles of V_c to this order. But even for very small V_{jk} 's it also has a solution $E_0(k)$ near E_k , given by

$$E_0(k) - E_k \cong \frac{(V_{jk})^2}{E_k - E_j} = (\delta E)_k. \quad (93)$$

This solution is closer to E_k , the weaker the coupling with j .

Let us now compute $\bar{a}_{jj}(E)$ at this energy, from (92). It is

$$\begin{aligned} \bar{a}_{jj}(E_0(k)) &\cong \lim_{\sigma \rightarrow 0} \left\{ 1 - \frac{1}{\sigma} \operatorname{Im} \left[\frac{(V_{jk})^2}{(\delta E)_k + i\sigma} \right] \right\}^{-1} \\ &= \frac{|V_{jk}|^2}{(V_{jk})^2 + (E_j - E_k)^2} \ll 1. \end{aligned} \quad (94)$$

In a similar way, one can show that

$$(j|F|k) \cong (E_j - E_k) / |V_{jk}|, \quad (95)$$

$$\lim_{V_{jk} \rightarrow 0} \bar{a}_{jk}(E_0(k)) = 0$$

so that

also.

In the finite random lattice with all orders of perturbations the same considerations obviously hold; now, however, E_k is a perturbed energy, and V_{jk} includes virtual effects. In principle all the same considerations apply, since except for the exact localized state E_k of (93) we have proved that all contributions to V 's and F 's fall off with distance sufficiently fast. Taking the limit $N \rightarrow \infty$ cannot change the above arguments for most states. However, because one of the distant wave functions may pop up at any point on the energy axis, it is not easy to see a way to assign a one-to-one correspondence of sites j and perturbed energies. This seems hardly necessary for a qualitative understanding of what is happening, however.

The difficulty lies in the fact that V_c is not really a continuous function in any sense. What can be done is to eliminate distant neighbors beyond a certain radius R . Then we find an appropriate perturbed energy $E_j'(R)$. The contribution to V_c , $\delta V_c(R)$, from beyond R is a probability variable which can be made to have as narrow a distribution as desired, but at all R may be large at a few E . Thus it is probable, but not certain, that a state localized around j has an energy within a predetermined $\delta V_c(R)$ of $E_j'(R)$. However, as we increase the size of the system we can never find an R beyond which $E_j'(R)$ is always as close as we like to the correct value. The fault lies in the Brillouin-Wigner technique, and is similar to problems as yet unsolved in other theories using this technique.

VI. ACKNOWLEDGMENTS

Aside from the help of Dr. L. R. Walker and Dr. P. A. Wolff already mentioned, I should like to thank a number of my colleagues for discussing this work with me; in particular, Dr. E. Abrahams, Dr. G. Feher, Dr. C. Herring, Dr. D. Pines, and Dr. G. H. Wannier.