Effect of Bardeen-Herring Correlation on Vacancy Diffusion in Anisotropic Crystals*

James G. Mullen

Argonne National Laboratory, Argonne, Illinois

(Received August 9, 1961)

From the Einstein relation, the principal diffusivities in anisotropic crystals are related to the atomic displacements. The effect of correlation is expressed in terms of correlation functions which are analogous to the correlation factor in isotropic diffusion. These correlation functions are calculated for vacancy self-diffusion in the primitive tetragonal, body-centered tetragonal, and hexagonal close-packed lattices by a computer technique. The possible use of correlation to determine mechanisms of anisotropic self-diffusion is discussed.

I. INTRODUCTION

SINCE Bardeen and Herring[1] noted that tracer diffusion is not a purely random walk process for certain mechanisms of diffusion, considerable theoretical and experimental progress has been made in calculating and measuring effects which are a consequence of correlated tracer diffusion. Theoretically, a number of transport properties should be altered as a result of correlation, e.g., deviations from the Einstein-Nernst relation[2] in ionic crystals, a reduction in the isotope effect[3] of diffusion, and deviations from an Arrhenius temperature dependence[4-6] in impurity diffusion. Only the first two of these effects have thus far been demonstrated experimentally.[7-11] The observation of the third effect will probably require considerable experimental accuracy in the absolute measurement of the temperature dependence of the diffusivity, since the primary effect of correlation is apparently to "shift" the activation energy,[6,10] while deviations from the Arrhenius relation appear to be small over the temperature range over which tracer diffusion measurements are typically made.

In all of the theoretical discussions to date dealing with the problem of impurity diffusion via vacancies, it has been necessary to set up a model of three or four frequency factors which are considered to deviate from the self-diffusion jump frequencies of a vacancy in the host lattice. These models are highly arbitrary in that vacancy jumps of different symmetry are equated and only vacancy jumps which either start or arrive at a nearest neighbor site relative to an impurity are con-

sidered appreciably different from the self-diffusion jump frequency. Long-range screening effects of the type discussed by Friedel[12-14] would appear to indicate that the number of vacancy jump frequencies which are significantly affected by a nearby impurity is considerably greater than the three or four used in impurity diffusion models.

Like impurity diffusion, anisotropic self-diffusion will manifest temperature-dependent correlation effects, since more than one vacancy jump frequency will govern diffusion. If each lattice site is a center of symmetry, e.g., zinc or tin, then it is possible to set up a model for vacancy diffusion which has the advantages of both being realistic and permitting a theoretical analysis of the contribution from correlation to diffusion. To date there has been but one attempt to determine the effect of correlation on anisotropic self-diffusion. Compaan and Haven[15] have treated a case with very special symmetry, for seven values of the relative frequency factors in and out of the basal plane, for Al sites in the corundum lattice. Their procedure involves setting up an electrical analog for each set of frequency factors and determining from appropriate measurements on the electrical network the various correlation factors. Because of the size of the networks involved, it is at best tedious to use this procedure for any realistic case with a sufficient number of relative frequency factors to be useful in analyzing experimental data.

In the present paper, the problem of correlation in anisotropic diffusion is first studied for a general lattice. This results in the introduction of "correlation functions" whose values lie between zero and unity. These correlation functions are explicitly calculated for the primitive tetragonal (pt), body-centered tetragonal (bct), and the hexagonal close-packed (hcp) lattices by a computer technique. For the bct and hcp lattices, the mean cosine of the angle between consecutive jumps does not form a simple power series, and the calculation

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

[1] J. Bardeen and C. Herring, *Imperfections in Nearly Perfect Crystals* (John Wiley & Sons, Inc., New York, 1952), p. 261.

[2] C. W. McCombie and A. B. Lidiard, Phys. Rev. 101, 1210 (1956).

[3] K. Tharmalingam and A. B. Lidiard, Phil. Mag. 44, 899 (1955).

[4] A. B. Lidiard, Phil. Mag. 46, 1218 (1955).

[5] A. D. Le Claire and A. B. Lidiard, Phil. Mag. 1, 518 (1956).

[6] J. R. Manning, Phys. Rev. Letters 1, 365 (1958).

[7] W. D. Compton and R. J. Maurer, J. Phys. Chem. Solids 1, 191 (1956).

[8] R. J. Friauf, Phys. Rev. 105, 843 (1957).

[9] A. S. Miller and R. J. Maurer, J. Phys. Chem. Solids 4, 196 (1958).

[10] J. G. Mullen, Phys. Rev. 121, 1649 (1961).

[11] A. H. Schoen, Phys. Rev. Letters 1, 138 (1958).

[12] J. Friedel, in *Advances in Physics*, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1954), Vol. 3, p. 446.

[13] J. Friedel, Suppl. Nuovo cimento 7, 287 (1958).

[14] J. S. Langer and S. H. Vosko, J. Phys. Chem. Solids 12, 196 (1959).

[15] K. Compaan and Y. Haven, Trans. Faraday Soc. 52, 786 (1961).

of correlation functions for these cases is consequently more involved, requiring additional approximations not required in the pt case or the case treated by Compaan and Haven.

II. EFFECT OF CORRELATION ON CRYSTALLINE DIFFUSION

It was shown by Einstein[16] that the diffusivity of a particle undergoing a sequence of random displacements is proportional to the mean-square displacement of each particle after a time τ. We briefly restate Einstein's derivation in the three-dimensional case to clarify notation and to emphasize the applicability of the derivation to both the correlated and uncorrelated random walk problems.

We assume the following: (a) Diffusion occurs in a system where the "marked" or tracer atoms are in a negligible concentration gradient, and (b) a time τ can be chosen sufficiently large that the probability of a tracer displacement of X, Y, Z, after the interval τ or $P(X,Y,Z,\tau)$, is symmetric to inversion in X, Y, Z, and at the same time small enough so that the initial distribution of tracer atoms $c(x,y,z,t)$ does not change appreciably in the interval τ. From the definition of $P(X,Y,Z,\tau)$, which from assumption (a) is independent of x, y, and z, we can write

$$c(x,y,z,\,t+\tau) = \int_{\text{all space}} c(x+X,\,y+Y,\,z+Z,\,t)$$
$$\times P(X,Y,Z,\tau)dXdYdZ. \quad (1)$$

Expanding the left side in powers of τ and the right side in powers of X, Y, Z, and using assumption (b) to discard higher order terms gives

$$c(x,y,z,t) + \tau\frac{\partial c(x,y,z,t)}{\partial t}$$
$$= c(x,y,z,t) + \frac{1}{2}\int\left(X\frac{\partial}{\partial x} + Y\frac{\partial}{\partial y} + Z\frac{\partial}{\partial z}\right)^2$$
$$\times c(x,y,z,t)P(X,Y,Z,\tau)dXdYdZ, \quad (2)$$

where we have made use of the relation $\int PdV = 1$, and $P(X,Y,Z,\tau) = P(-X, -Y, -Z, \tau)$.

Equation (2) is equivalent to Fick's phenomenological diffusion equation,

$$\frac{\partial c}{\partial t} = \sum_{l,m=1}^{3} D_{lm}\frac{\partial^2 c}{\partial x^l \partial x^m}, \quad (3)$$

provided we make the identification

$$D_{lm} = \frac{1}{2\tau}\int_{\text{all space}} X^l X^m P(X^1,X^2,X^3,\tau)dX^1 dX^2 dX^3. \quad (4)$$

---

[16] See, for example, A. Einstein, *Theory of Brownian Motion* (Dover Publishers, New York, 1956), p. 12.

The symbols X^1, X^2, X^3 and x^1, x^2, x^3 are used interchangeably with X, Y, Z, and x, y, z, respectively. Since the symmetry of $P(X,Y,Z,\tau)$ is unaltered whether the diffusion mechanism is correlated or not, Eq. (4) is applicable to both correlated and uncorrelated diffusion. Also for those cases where P is symmetric to reflection in each of the X^i, in addition to inversion about the origin, then all of the $D_{lm} = 0$, for $l \neq m$. All of the lattices considered subsequently will have sufficient symmetry that such a principal set of axes can be determined by inspection. Having a principal set of axes the principal diffusivities are

$$D_{xx} = \langle X^2\rangle/2\tau, \quad D_{yy} = \langle Y^2\rangle/2\tau, \quad D_{zz} = \langle Z^2\rangle/2\tau, \quad (5)$$

where the notation $\langle \ \rangle$ stands for an average defined by Eq. (4). For the isotropic case $\langle X^2\rangle = \langle Y^2\rangle = \langle Z^2\rangle = \frac{1}{3}\langle R^2\rangle$, where $\langle R^2\rangle$ is the total mean displacement of a tracer after a time τ. Hence, for the isotropic case, $D = (1/6\tau) \times \langle R^2\rangle$. This relation which has been used as a basis for previous calculations of the effects of correlation on diffusion is not meaningful in the anisotropic case, and the theory must be formulated using Eq. (5) as a basis.

The Einstein relation, represented by Eq. (5), can also be expressed in terms of the individual atomic displacements. If x_i represents the projection of the ith tracer jump along the x axis, then

$$D_{xx} = \frac{1}{2\tau}\langle(\sum_{i=1}^{n} x_i)^2\rangle, \quad (6)$$

with similar expressions obtaining for D_{yy} and D_{zz}. The average in Eq. (6) is obtained by taking the weighted mean of all of the possible $(\sum x_i)^2$ in the interval τ. When diffusion is correlated, the cross terms, $\langle x_i x_j\rangle$, $i \neq j$, in Eq. (6) do not vanish. In vacancy diffusion the cross terms do not average to zero because of the asymmetry in the vacancy distribution about a tracer following an initial vacancy-tracer exchange. The mean square displacement along the principal axes can be rewritten as

$$\langle X^2\rangle = \sum_{i=1}^{n} \langle x_i^2\rangle + 2\langle x_1 x_2 + x_1 x_3 + \cdots x_1 x_n\rangle$$
$$+ 2\langle x_2 x_3 + \cdots x_2 x_n\rangle + \cdots 2\langle x_{n-1} x_n\rangle. \quad (7)$$

In evaluating the cross terms for vacancy diffusion it will be assumed that a tracer which exchanges with a vacancy will not make any exchanges with a different vacancy until the initial vacancy is far from the tracer. Because there is a significant probability that a vacancy will exchange with a tracer even after 20 or more vacancy jumps following an initial vacancy-tracer exchange, correlation factors derived using this assumption could be significantly in error for systems where vacancy concentrations are large. (In cases where this assumption is not valid, correlation factors will be temperature dependent even for isotropic self-diffusion.) Also, we will assume that $\langle x_i x_{i+j}\rangle$ is negligibly small for large j, and that n is sufficiently large that each of the

$n-1$ sets of cross terms shown in Eq. (7) can be considered to have an infinite number of terms, i.e., we replace

$$\sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \langle x_i x_{i+j} \rangle$$

by

$$\sum_{i=1}^{n-1} \sum_{j=1}^{\infty} \langle x_i x_{i+j} \rangle.$$

For diffusion in single crystals, many of the $n-1$ sets of cross terms will be equivalent, i.e., even for anisotropic materials many of the tracer displacements are equivalent in the sense that

$$\sum_{j=1}^{\infty} \langle x_i x_{i+j} \rangle$$

is the same number for a large fraction of the $n-1$ values of i. If there are N sets of displacements which are equivalent, and the αth set occurs n_α times in the interval τ, then Eq. (7) will take the form

$$2\tau D_{xx} = \sum_{\alpha=1}^{N} n_\alpha x_\alpha^2 f_{\alpha x},$$

where

$$f_{\alpha x} = 1 + 2 \sum_{j=1}^{\infty} \xi_j^\alpha, \tag{8}$$

and

$$\xi_j^\alpha = \langle x_\alpha x_{\alpha+j} \rangle / x_\alpha^2,$$

with similar expressions for $\langle Y^2 \rangle$ and $\langle Z^2 \rangle$, with x replaced by y and z, and ξ by ψ and ω, respectively. The effect of correlation is then accounted for by determining $f_{\alpha x}$, $f_{\alpha y}$, and $f_{\alpha z}$ for all α. This calculation requires a specification of the lattice and a model of the vacancy frequency factors in the vicinity of a tracer.

III. SPECIAL CASES

A. Isotropic Diffusion

We shall now calculate the effect of correlation on diffusion in a two-dimensional square lattice. This is done to illustrate the computer technique used in the subsequent anisotropic problems, for a case which is easily visualized and physically simpler than anisotropic diffusion. The geometry of this lattice is shown in Fig. 1. We consider a tracer initially at the site labeled 1. After a period of time the tracer will exchange with a vacancy and move to an adjacent site labeled 2. For the isotropic lattices there is only one value of α since only one type of jump contributes to the diffusion along any principle axis. For the two-dimensional square lattice, Eq. (8) takes the form $D_{xx} = (1/2\tau)(n/2)a^2 f$, where $f = 1 + 2 \sum_1^\infty \xi_j$. The subscript $\alpha = 1$ has been omitted to simplify notation. n is the total number of tracer jumps after a time τ, of which only half contribute to diffusion along the x axis.
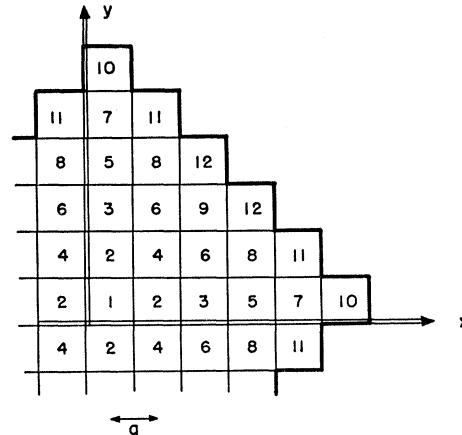

FIG. 1. Two-dimensional square lattice. The heavy border line shows finite boundary approximation and labeling of lattice sites.

The isotropic problem consists of essentially two parts; first, the calculation of ξ_1, or the difference in the "forward" and "backward" jump probability of a tracer, given an initial vacancy-tracer exchange, and second, expressing ξ_j as a simple power of ξ_1 through the relation $\xi_j = \xi_1^j$, which permits the exact summation of the expression for f to be performed.

For the two-dimensional square lattice (Fig. 1), ξ_1 is given by

$$\xi_1 = W_3 - W_1,$$

where W_3 is the probability that the tracer jump following the initial jump is to site 3, and W_1 is the probability that this second jump is to site 1. To calculate $W_3 - W_1$ we must examine the diffusion of the vacancy from site 1. If we designate $p_n(k)$ as the probability that the nth lattice site of a crystal is occupied after k jumps by a vacancy which is considered to start at the lattice site 1, then the sum of the probabilities that the lattice site n is visited after an infinite number of jumps is

$$P_n = \sum_{k=0}^{\infty} p_n(k).$$

This sum of probability is simply related to $W_3 - W_1$, i.e.,

$$W_3 - W_1 = A(P_3 - P_1),$$

where A is the probability that a vacancy at site 3, or at site 1, will move to the particular site 2 where the tracer resides after the initial exchange, and is equal to $\frac{1}{4}$ for the two-dimensional square lattice. The distribution of probability after k vacancy jumps is related to the distribution after $k-1$ jumps by a linear set of equations

$$p_{n'}(k) = \sum_n T_{n'n} p_n(k-1), \tag{9}$$

where $T_{n'n}$ is the probability of a transition from the site n to the site n'. The number of equations will, in general, equal the number of lattice sites in the crystal. Since the vacancy is considered to start at site 1, the
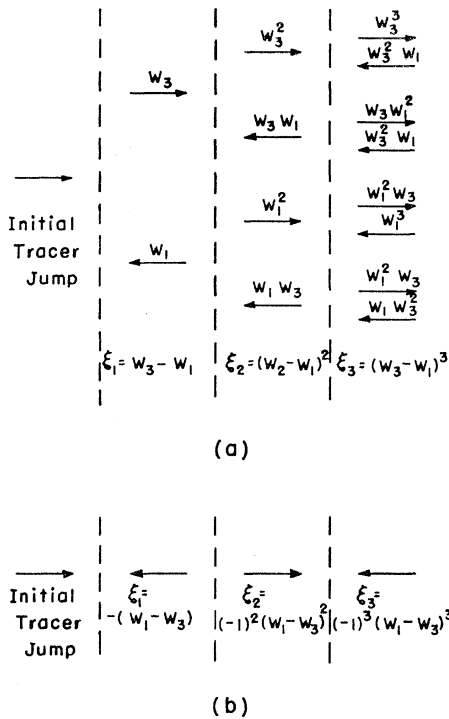
(a)

(b)

FIG. 2. (a) The possible sequences of tracer jumps along the x axis due to a single vacancy with the appropriate probabilities for each sequence. (b) Net probabilities for successive "backward" jumps for $k=1, 2, 3$, with the coefficient $(-1)^k$ indicating whether the sense of the sequence is the same or opposite to the initial jump.

initial probability distribution can be written as a column matrix

$$\mathbf{p}(0) = \{1, 0, \cdots, 0\}.$$

Expressing the set of linear equations (9) in matrix notation, the sum of the probabilities at each lattice site is

$$\mathbf{P} = \sum_{k=0}^{\infty} \mathbf{p}(k) = \sum_{k=1}^{\infty} (1 + \mathbf{T}^k) \mathbf{p}(0). \qquad (10)$$

Performing the summation on k gives

$$\mathbf{P} = (1 - \mathbf{T})^{-1} \mathbf{p}(0). \qquad (11)$$

The Bardeen-Herring[1] approach to the problem of evaluating P_1 and P_3 is to set up a matrix diagram, and by directly following the diffusion of probability to evaluate the contributions $\mathbf{p}(0)(1 + \mathbf{T} + \mathbf{T}^2 + \cdots)$. This procedure does not lead to a rapidly converging series, although Manning[17] has extended this approach to accurately calculate correlation factors in isotropic crystals. The approach to the problem used here is to construct a symmetric, imaginary boundary about the site 1, and to assume that vacancies which get out of this boundary can be considered as making a negligible contribution to the difference $P_3 - P_1$. This finite boundary restricts the size of the matrix $\mathbf{T}$, and permits

the direct evaluation of Eq. (11). The way in which P_1 and P_3 have been defined requires that transitions to site 2 be forbidden. This restriction need not be adhered to, however, for cases where each lattice site is a center of symmetry, as will be the case in all subsequent examples, since we are only interested in the difference $P_3 - P_1$, which is unaffected by transitions to site 2. Relaxing this restriction greatly reduces the size of the matrix $\mathbf{T}$ since many of the sites will now be equivalent by symmetry. For the boundary chosen in Fig. 1, involving 96 lattice sites, only 12 are unique and require different labels.

For the purpose of illustration we write the matrix $\mathbf{T}$ which would obtain if only the first six sites shown in Fig. 1 were used as a boundary. Within this approximation we have

$$\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}.$$

As for the second part of the problem, we can show that $\xi_j = \xi_1{}^j$ by considering the possible sequences of tracer steps and explicitly calculating the probability of each of these possibilities. The possible sequences of steps with the appropriate probability for each is shown schematically in Fig. 2(a) for k equal to 1, 2, and 3. By this diagrammatic procedure we are led directly to the relation $\xi_j = \xi_1{}^j$. An even simpler way to arrive at this result is to regard ξ_1 as (-1) times the net probability of a "backward" jump. The relation $\xi_j = \xi_1{}^j$ is then obvious, since the probability of a particular sequence of events is a product of the individual probabilities. This solution is diagrammatically depicted in Fig. 2(b). We may now sum the geometrical series for f in Eq. (8), with the result

$$f = (1 + \xi_1)/(1 - \xi_1).$$

The size of the matrix $\mathbf{T}$ is, of course, determined by the accuracy desired in a given calculation. By examining specific examples, it has been found that a boundary about twice the distance from the site 1 to the outermost site relevant to ξ_1 (3 in this case), will yield very accurate values of f. For the example of a square lattice, if 9 sites are considered in $\mathbf{T}$ the resulting f is 0.4679; for 12 sites the result is 0.4674. By an electrical analog technique, Compaan and Haven[18] have found $f = 0.4669$. Hence the error in the calculation is only 0.1%, when a quite manageable $\mathbf{T}$ is used with the above procedure.

The above analysis can be used for all of the isotropic lattices. In cases tried, excellent agreement exists with the second set of values tabulated by Compaan and Haven. For example, using a 34×34 T matrix

[17] J. R. Manning, Phys. Rev. 116, 819 (1959).

[18] K. Compaan and Y. Haven, Trans. Faraday Soc. 54, 1498 (1958).

for the cubic lattice gave a value of 0.6533 by the present technique, compared with 0.6531 given by Compaan and Haven.

It should be emphasized that while the present procedure can be used for isotropic diffusion, its real advantage is for the anisotropic cases where one desires a systematic variation in the various vacancy jump frequencies.

B. Tetragonal Lattice

The problem of correlation in anisotropic lattices generally has two additional complications not found in the isotropic case. First, there are generally more than one frequency factor involved, which means that the matrix $\mathbf{T}$, which depends on the relative frequency factors, is not constant. Second, ξ_j, ψ_j, and ω_j cannot generally be expressed as a simple power of ξ_1, ψ_1, and ω_1. In the tetragonal lattice, vacancy diffusion is particularly simple in that only the first difficulty is encountered. Hence, by the same argument used for the two-dimensional square lattice, we can show

$$\xi_j=\xi_1{}^j \quad \text{and} \quad \omega_j=\omega_1{}^j. \tag{12}$$

The geometry of the tetragonal lattice and the boundary used to determine the matrix $\mathbf{T}$, with appropriate labeling, are shown in Fig. 3. As the figure indicates, ν_A is the mean frequency for vacancy jumps in the basal plane and ν_B is the mean frequency for vacancy jumps out of the basal plane, or along the c axis of the lattice.

The connection between the principal diffusivities and the correlation functions can be written from an inspection of Eq. (8). If we let p be the probability that a tracer has a vacancy at an adjacent site, either in the
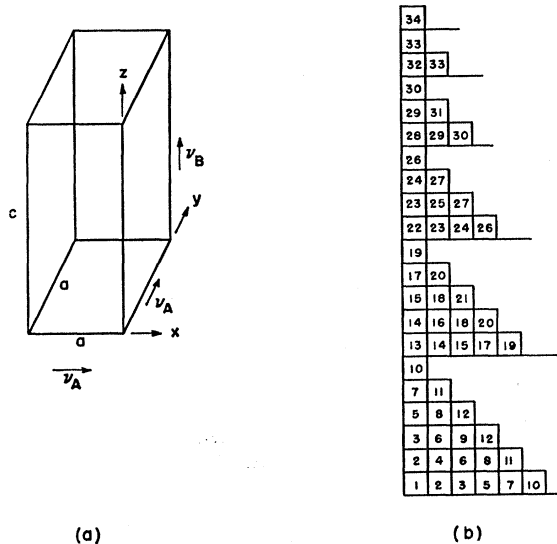

(a)

(b)

FIG. 3. (a) Geometry of the primitive tetragonal lattice. (b) Finite boundary and labeling used to specify matrix $\mathbf{T}$. Each triangular region from bottom to top represents subsequent layers of the tetragonal lattice along the c axis.
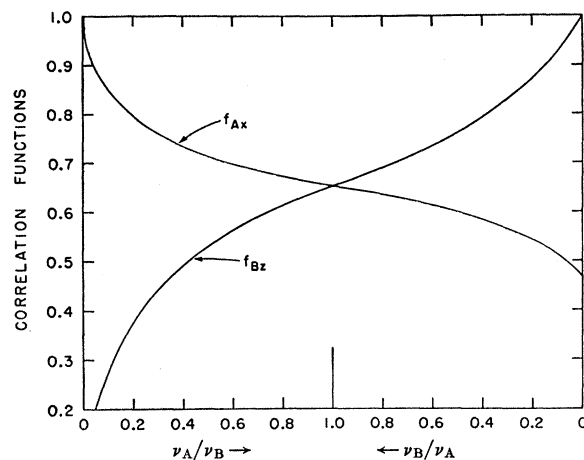

FIG. 4. Correlation functions for the primitive tetragonal lattice as a function of the ratio of frequency factors.

basal plane or along the c axis, then

$$n_A=\tfrac{2}{3}p\tau\nu_A,$$
and
$$n_B=\tfrac{1}{3}p\tau\nu_B, \tag{13}$$

since there are twice as many sites adjacent to a tracer in the basal plane as there are along the c axis. Combining Eq. (13) with Eq. (8) gives

$$D_{xx}=\tfrac{1}{6}\nu_A p a^2 f_{Ax},$$
$$D_{zz}=\tfrac{1}{6}\nu_B p c^2 f_{Bz}, \tag{14}$$
and
$$(c/a)^2 D_{xx}/D_{zz}=\nu_A f_{Ax}/\nu_B f_{Bz}.$$

Because of the applicability of Eq. (12), f_{Ax} and f_{Bz} can be expressed in terms of ξ_1 and ω_1, i.e.,

$$f_{Ax}=(1+\xi_1)/(1-\xi_1), \quad \text{and} \quad f_{Bz}=(1+\omega_1)/(1-\omega_1). \tag{15}$$

In terms of the labeling shown in Fig. 3 we have $\xi_1 = (P_3-P_1)A$ and $\omega_1=(P_{22}-P_1)B$, where A and B are the probabilities that a vacancy will make an A- and B-type jump, respectively. In terms of the jump frequencies ν_A and ν_B, we have $A=\nu_A/(4\nu_A+2\nu_B)$ and $B=\nu_B/(4\nu_A+2\nu_B)$. The matrix elements P_1, P_3, and P_{22} were calculated from Eq. (11) by the same procedure used in the two-dimensional square lattice. For this case $\mathbf{T}$ is a 34×34 matrix whose elements vary as the relative frequency factors are varied. It is because of this systematic variation in $\mathbf{T}$ that the computer approach is so convenient. f_{Ax} and f_{Bz} were calculated, using 20 values of A/B ranging from 0 to 1 and 20 values of B/A in the same interval. The results of this calculation[19] are shown in Fig. 4. In the isotropic limits $A/B \to 0, 1, \infty$, we find $f_{Ax}=1.0000, 0.6533, 0.4674$, and $f_{Bz}=0, 0.6533, 1.0000$ in excellent agreement with the Compaan and Haven values corresponding to these limits. These values of f_{Ax} and f_{Bz} were used to calcu-

---

[19] Tables of the functions shown in Figs. 4, 5, 7, and 8 can be obtained from the author on request.
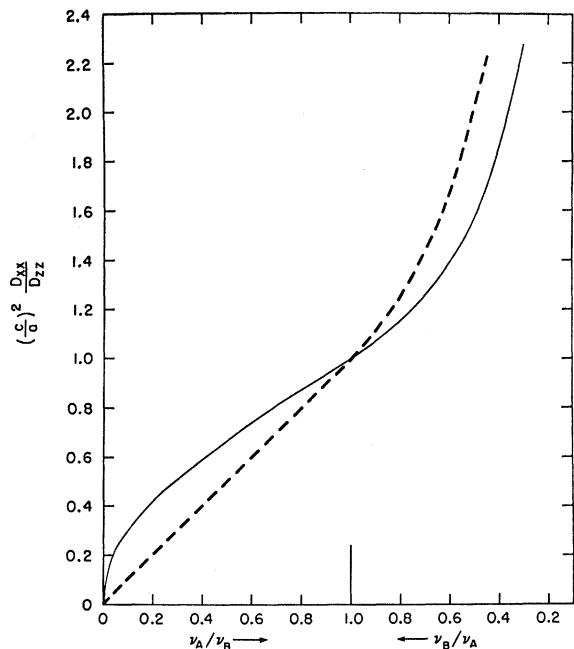
FIG. 5. Diffusion anisotropy for the primitive tetragonal lattice. The solid line obtains for vacancy diffusion, and the dashed line for uncorrelated diffusion with the same set of basis jump vectors.

late $(c/a)^2 D_{xx}/D_{zz}$, from Eq. (14), as a function of ν_A/ν_B. The result of this calculation is shown in Fig. 5. By comparing experimental values of $(c/a)^2 D_{xx}/D_{zz}$ with the function graphed in Fig. 5 we can determine ν_A/ν_B with the effect of correlation included. From ν_A/ν_B the correlation functions f_{Ax} and f_{Bz} are readily determined from the calculated values of these functions shown in Fig. 4.

C. Body-Centered Tetragonal and Hexagonal Close-Packed Lattices

For the bct and hcp lattices, we shall carry out an analysis similar to that found in the previous section. The geometry of these lattices is shown schematically in Figs. 6(a) and 6(b), with ν_A representing the mean frequency that a vacancy jumps in the basal plane and ν_B the frequency of jumps out of the basal plane. The connection between the principal diffusivities and the correlation functions follows directly from Eq. (8). Using arguments completely analogous to the tetragonal case, it can be shown that

$$D_{xx}=\tfrac{1}{12}pa^2(\nu_A f_{Ax}+\nu_B f_{Bx}), \qquad \text{(bct)}$$
$$D_{zz}=\tfrac{1}{12}pc^2\nu_B f_{Bz}, \qquad \text{(bct)} \quad (16)$$
$$(c/a)^2 D_{xx}/D_{zz}= (1/f_{Bz})(\nu_A f_{Ax}/\nu_B+f_{Bx}), \quad \text{(bct)}$$

and

$$D_{xx}= (1/24)pa^2(3\nu_A f_{Ax}+\nu_B f_{Bx}), \quad \text{(hcp)}$$
$$D_{zz}=\tfrac{1}{16}pc^2\nu_B f_{Bz}, \qquad \text{(hcp)} \quad (17)$$
$$(c/a)^2 D_{xx}/D_{zz}= (2/f_{Bz})(\nu_A f_{Ax}/\nu_B+\tfrac{1}{3}f_{Bx}). \quad \text{(hcp)}$$

To calculate the correlation functions for the hcp and bct cases requires an alteration in the analysis previously used for the tetragonal case, since expressions analogous to Eq. (12) do not obtain for ξ_j. For these cases the procedure used to express $\xi_j (j>1)$ in terms of ξ_1 is as follows: We partition ξ_j into two parts, a major part which can be expressed in a form similar to Eq. (12), and a lesser part which we will calculate explicitly by considering the contribution from the remaining possible tracer jumps following an initial tracer-vacancy exchange. We proceed by calculating ξ_1^A and ξ_2^A explicitly, where the superscript A indicates that we refer only to the case where the initial vacancy-tracer exchange is of the A type, i.e., where $\alpha=A$. Following the initial exchange, the next tracer jump can be either of the A or B type, or symbolically the possible sequences can be written (AA) or (AB). Thus we can write

$$\xi_1^A=\xi_1(AA)+\xi_1(AB),$$

where $\xi_1(AA)$ is the contribution to ξ_1^A from (AA) jumps and $\xi_1(AB)$ is that from (AB) jumps. Considering two jumps following the initial exchange, we have for the possible sequences (AAA), (AAB), (ABB), and (ABA), so that

$$\xi_2^A=\xi_2(AAA)+\xi_2(AAB)+\xi_2(ABB)+\xi_2(ABA).$$

By the same type of argument used to arrive at the relation $\xi_j=\xi_1^j$ for the two-dimensional square lattice we can show that $\xi_2(AAA)=\xi_1(AA)\xi_1(AA)$, $\xi_2(AAB)=\xi_1(AA)\xi_1(AB)$, etc. Thus

$$\xi_2^A=\xi_1(AA)\xi_1(AA)+\xi_1(AA)\xi_1(AB)$$
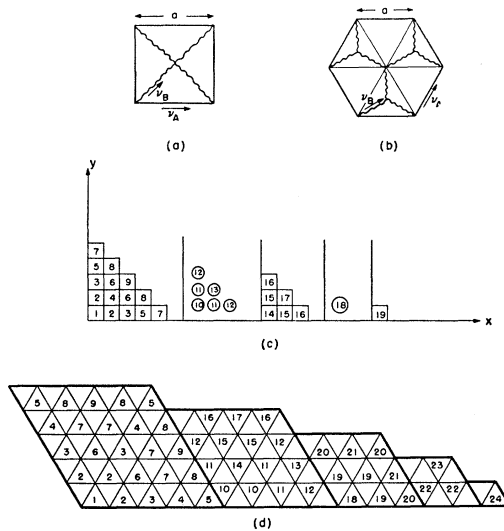$$+\xi_1(AB)\xi_1(BB)+\xi_1(AB)\xi_1(BA). \quad (18)$$

FIG. 6. (a) and (b) Schematic representation of the lattice geometry of the bct and hcp lattices, respectively, as viewed along the c axis. (c) and (d) Finite boundary and labeling used to specify matrix **T** for bct and hcp lattices, respectively. Successive units represent layers along the c axis.

TABLE I. Dependence of ξ_1 and ω_1 on elements of matrix **P**.

	bct	hcp
$\xi_1(AA)$	$A(P_3-P_1)$	$A(P_3+P_6-P_2-P_1)$
$\xi_1(AB)$	$2B(P_{11}-P_{10})$	$B(P_{11}-P_{10})$
$\xi_1(BB)$	$B(P_4+P_{17}-P_1-P_{14})$	$B(P_2+P_{19}-P_1-P_{18})$
$\xi_1(BA)$	$2A(P_{11}-P_{10})$	$3A(P_{11}-P_{10})$
$\omega_1(BB)$	$B(P_{14}+2P_{15}+P_{17}-P_1-2P_2-P_4)$	$B(P_{18}+2P_{19}-P_1-2P_2)$

A similar expression can be written for $\xi_j{}^B$. We find that in general

$$\xi_j{}^A=\xi_1(AA)^j+\text{cross terms,}$$
$$\xi_j{}^B=\xi_1(BB)^j+\text{cross terms.} \tag{19}$$

Physically, it would be expected that the cross terms make a small contribution to the correlation functions in Eqs. (16) and (17), since (AB) and (BA) jump sequences require a minimum of two jumps of a vacancy to one of the tracer. We can show this explicitly by calculating $\xi_1(AB)$ and $\xi_1(BA)$ by the matrix technique previously discussed. The ξ's and ω's are obtained from inspection of the lattice diagram shown in Figs. 6(c) and 6(d). Their dependence on the matrix elements is shown in Table I for the bct and hcp lattices. Like the pt case, A and B are the probabilities of vacancy jumps in and out of the basal plane, respectively. Their relation to the jump frequencies is $A=\nu_A/(4\nu_A+8\nu_B)$ and $B=\nu_B/(4\nu_A+8\nu_B)$ for the bct lattice, and $A=\nu_A/(6\nu_A+6\nu_B)$ and $B=\nu_B/(6\nu_A+6\nu_B)$ for the hcp lattice. The elements of the matrix **P** shown in Table I were determined from Eq. (11), as before, with **T** specified by the boundaries and notation shown in Figs. 6(c) and
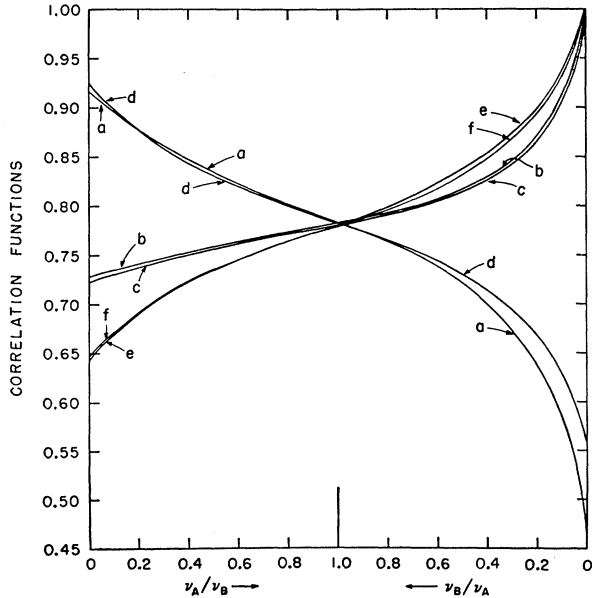

FIG. 7. Correlation functions for the bct lattice; (a) $f_{Ax}(\nu_A/\nu_B)$, (b) $f_{Bx}(\nu_A/\nu_B)$, and (c) $f_{Bz}(\nu_A/\nu_B)$. Correlation functions for the hcp lattice; (d) $f_{Ax}(\nu_A/\nu_B)$, (e) $f_{Bx}(\nu_A/\nu_B)$, and (f) $f_{Bz}(\nu_A/\nu_B)$.
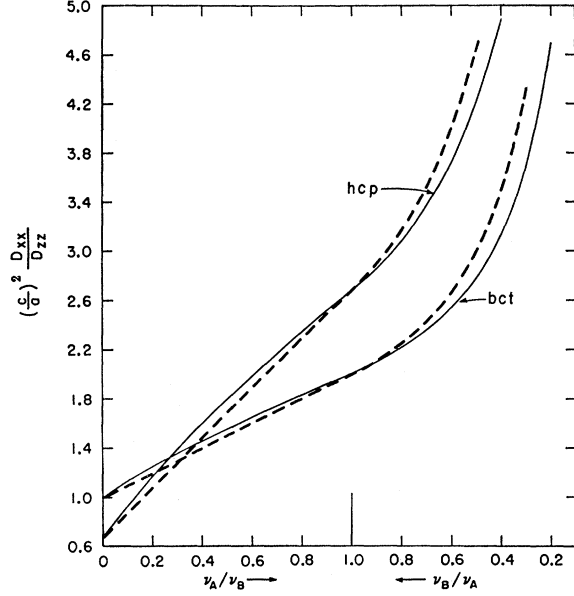

FIG. 8. Diffusion anisotropy for the bct and hcp lattices. The solid lines obtain for vacancy diffusion and the dashed lines for uncorrelated diffusion with the same set of basis jump vectors.

6(d), for the bct and hcp cases, respectively. As a result of a numerical calculation, it was found that for the hcp case the largest value of $\xi_1(AB)$ was 0.048 which occurred in the limit $A/B \to 0$, and the largest value of $\xi_1(BA)$ was 0.038 in the limit $A/B \to 1.43$. Since the cross terms contribute less than 10% to the correlation functions in first order, it will be sufficient to account for these terms to second order with a resulting error considerably less than 1% over the entire range of ν_A and ν_B. The contribution from the cross terms is of a similar magnitude for the bct case.

Summing the contributions to the correlation functions which form a geometrical series exactly, and the cross terms to second order, gives

$$f_{Ax}=\frac{1+\xi_1(AA)}{1-\xi_1(AA)}+2\xi_1(AB)$$

$$\times[1+\xi_1(AA)+\xi_1(BB)+\xi_1(BA)],$$

$$f_{Bx}=\frac{1+\xi_1(BB)}{1-\xi_1(BB)}+2\xi_1(BA) \tag{20}$$

$$\times[1+\xi_1(BB)+\xi_1(AA)+\xi_1(AB)],$$

$$f_{Bz}=\frac{1+\omega_1(BB)}{1-\omega_1(BB)}.$$

This relation applies to both the bct and hcp lattices, provided the appropriate ξ's and ω's shown in Table I are used. The results of the numerical calculations of f_{Ax}, f_{Bx}, and f_{Bz} over the entire range of ν_A/ν_B, with the same intervals as for the pt case, are shown in Fig. 7.

The diffusion anisotropy as calculated from Eqs. (16) and (17) is shown in Fig. 8.

Since the accuracy of the present calculation depends on the size of the matrix $\mathbf{T}$, i.e., the size of the region chosen in the finite boundary approximation, the correlation functions calculated for hcp and bct lattices are somewhat in error. This error is estimated to be of the order of 1% or less over the entire range of ν_A/ν_B. For the bct lattice, f_{Bx} and f_{Bz} differ by about 1% in the limit $\nu_A/\nu_B \to 0$, which is probably because the matrix element P_{17} used in f_{Bz} is so near the edge of the chosen boundary. Because of present-day experimental limitations and the increase in computer time required for a larger $\mathbf{T}$ matrix, a more accurate calculation was not believed to be warranted.

As can be seen from Fig. 7, f_{Bx} and f_{Bz} are very nearly the same over the entire range of ν_A/ν_B. Since the calculated differences between f_{Bx} and f_{Bz} are of the same order as the accuracy of the calculation, these differences must not be considered quantitative. The fact that f_{Bx} and f_{Bz} are the same to a good approximation, however, shows that the effect of correlation for the hcp and bct lattices can be considered as primarily resulting in a scaling down of A- and B-type jumps with scaling factors f_A and f_B. It can also be seen from Fig. 7 that the magnitude of the correlation functions changes rapidly in the limits $A/B \to 0$ and $B/A \to 0$. This indicates that even for highly anisotropic crystals where one frequency factor is much larger than the other, the correlation functions may be considerably different from those expected using a one-frequency model. For example, if B jumps occur only 5% as frequently as A jumps in an hcp lattice, then f_{Ax} is altered from 0.560 to 0.605, or about 8%, from the value obtained neglecting B jumps.

IV. DISCUSSION

Recent data on anisotropic self-diffusion are not extensive. Shirn, Wajda, and Huntington[20] have shown that considerable information can be inferred about diffusion mechanisms in anisotropic materials from measurements of the temperature dependence of the diffusion anisotropy. Their method of analysis, however, does not permit one to distinguish between mechanisms which have the same set of basis jump vectors such as the vacancy and exchange mechanisms, or more generally the ring mechanism. In principle, these mechanisms could be differentiated by sufficiently precise data, for the anisotropic lattices previously discussed. For ionic crystals, departures from the Einstein-Nernst relations could be used to distinguish between certain diffusion mechanisms. The distinction between

correlated and uncorrelated mechanisms could be established by determining the ratio of the principal diffusivities to the principal conductivities, since this ratio is a constant for the uncorrelated mechanism but not for the correlated mechanism. For metals it should be possible to determine mechanisms of diffusion by accurate measurements of deviations from a Boltzmann temperature dependence for such quantities as D_{zz}. A quantitative measurement of this effect which is in accord with that predicted in the preceding analysis would constitute strong evidence for the vacancy mechanism of diffusion.

Of the metallic systems for which the temperature dependence of the principal diffusivities have so far been measured, either the temperature dependence of the anisotropy has been small or the data too imprecise to use the above criteria to determine mechanisms of self-diffusion. In this regard it should be noted that highly anisotropic materials might be the most favorable for observing these effects. For example, in cases where $D_{xx} \gg D_{zz}$, i.e., $\nu_A \gg \nu_B$, then f_{Bz} will be very sensitive to small changes in ν_B and the deviations from exponential in D_{zz} could be appreciable.

In cases where the vacancy mechanism is responsible for diffusion, it should be noted that both the magnitude and temperature dependence of the frequency factors inferred from anisotropic diffusion data can be appreciably in error when the effect of correlation is neglected. As an example of the error which can result from neglecting correlation, we have used the data of Shirn et al.[20] for zinc to calculate ν_A/ν_B as a function of temperature, both neglecting and including the effect of correlation. From the data of Shirn et al. we take $D_{xx}/D_{zz} = 0.75$ at $1000/T = 1.5(°\text{K})^{-1}$ and $D_{xx}/D_{zz} = 0.441$ at $1000/T = 2.0(°\text{K})^{-1}$. Assuming an exponential temperature dependence for the frequency factors as predicted by reaction rate theory,[21] we find

neglecting correlation:

$$\nu_A/\nu_B = 10.9 \exp(-3.22 \text{ kcal}/RT),$$

including correlation:

$$\nu_A/\nu_B = 17.0 \exp(-3.83 \text{ kcal}/RT).$$

It should be noted that in addition to the large error in pre-exponential factors, the calculated difference in energy of motion of A- and B-type jumps is 16% in error when correlation is neglected. This effect could be much larger in crystals showing a large temperature dependence to the diffusion anisotropy.

ACKNOWLEDGMENT

[20] G. A. Shirn, E. S. Wajda, and H. B. Huntington, Acta Met. 1, 513 (1953).

[21] G. H. Vineyard, J. Phys. Chem. Solids 3, 121 (1957).