ANNALS OF PHYSICS: **21**, 494–549 (1963)

Studies of Polaron Motion

Part III: The Hall Mobility of the Small Polaron*

L. FRIEDMAN AND T. HOLSTEIN

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania

The one-dimensional molecular crystal model of polaron motion, developed in parts I and II, is suitably generalized to consider the existence of a Hall effect. As in II, the treatment is confined to the case for which the electronic-overlap term of the total Hamiltonian is a small perturbation. In zeroth order—i.e., in the absence of this term—the electron is localized at a given site. The vibrational state of the system is specified by a set of quantum numbers, N_k, giving the degree of excitation of each vibrational mode. The existence of a nonvanishing electronic bandwidth then gives rise to transitions to neighboring sites.

Of principal interest in the present paper is the high temperature regime ($T > T_t$, as defined in II) where polaron motion is predominantly by means of random jumps between neighboring sites. Although the lowest order jump rate is adequate in considering the polaron drift mobility, higher order processes, involving the occupation of (at least) three sites, must be taken into account in treating the Hall effect. In particular, it is demonstrated the relative probability of the electron, initially located on one of the three sites, hopping to one or the other of the remaining two sites, is modified by a contribution which, both in sign and magnitude, is linearly proportional to the applied magnetic field. This effect is shown to arise from the interference between the amplitude for the direct transition from the initial to the final site, and the amplitude for an indirect, second order transition, involving intermediate occupancy of the third site.

The (magnetic) field induced components of the jump rates, corresponding to the above processes, are first calculated by a classical occurrence-probability approach which treats the lattice vibrational coordinates as given functions of time. The second approach presents a full quantum mechanical calculation of the jump rates. The results of this treatment agree with those of the occurrence-probability approach in the classical limit ($T \gg \Theta_{\text{Debye}}$).

The order of magnitude of the calculated Hall Coefficients are found to be greater than or comparable to the "normal" result ($R = -1/nec$) depending on whether or not the three sites involved in the transition are mutually nearest neighbors. A final note corrects an error in a previous paper by one of the authors (T. H.) concerning the sign of the Hall effect in impurity conduction.

* This work was supported by both the Westinghouse Electric Corporation and the U. S. Air Force Office of Scientific Research, under contract number AF49(638)-323.

INTRODUCTION

The present work constitutes an extension of the theory of the *small polaron* developed in two previous papers (*1*). As discussed in those papers, the concept of the small polaron is essential for the understanding of the transport properties of *narrow-band* semiconductors—specifically, semiconductors whose electronic conduction bandwidths (or, in the case of hole-transport, valence bandwidths) are $\lesssim 0.5$ ev. In such materials, the interaction between lattice vibrations and an excess charge carrier (electron or hole) is generally sufficiently strong so as to give rise to a situation in which the charge carrier is essentially self-trapped in the immediate neighborhood of any one of an infinite number of crystallographically equivalent atomic sites; the existence of nonvanishing electronic bandwidth manifests itself in occasional transfers among these sites. In accordance with the nomenclature of I and II, the unit consisting of the charge carrier and the surrounding induced lattice deformation—the latter providing the potential well within which the carrier is trapped—is designated as the *small polaron.*

Previously developed polaron treatments (*2*, *3*) based on the continuum-polarization model in which (a) the ionic dipole moments arising from the displacements of discrete lattice particles are replaced by a continuous polarization density, and (b) the motion of the excess charge carrier is formulated in terms of the effective mass approximation, are not adequate for the description of the small polaron; in a proper theory, the periodicity of the actually discrete lattice structure must be taken explicitly into account.

A basic feature of the theory contained in I and II is the prediction of two types of small-polaron motion prevalent at temperatures above and below a certain transition temperature T_t (which in practice is $\sim\frac{1}{2}\Theta_{\text{Debye}}$), respectively. In the low temperature regime ($T < T_t$) the motion is describable in terms of a band picture in which the width of the "polaron-band" is an exponentially decreasing function of temperature. In the high-temperature regime ($T > T_t$), the band picture is not applicable. The motion of the charge carrier consists of a succession of mutually incoherent thermally-activated jumps between neighboring sites, each of which is accompanied by multiple phonon emission (or absorption).

The above-described two mechanisms of polaron motion each give rise to a characteristic temperature variation of the polaron drift-mobility, μ_D. In the low-temperature regime, μ_D is a diminishing function of temperature (as is to be expected from the decrease of polaron bandwidth with increasing temperature). On the other hand, when $T > T_t$, μ_D is an exponentially rising function of temperature, characteristic of an activation-type transport process. A temperature variation of this latter type is believed (*4*, *5*) to occur in certain transition-metal oxides (notably NiO).

The elementary processes giving rise to the polaron drift mobility are the above-described transitions between two (nearest neighbor) sites. In contrast, the existence of a nonvanishing Hall effect (*6*) depends on processes involving the (coherent) occupation of at least three sites. In particular, it will be demonstrated in this paper that the relative probabilities of an electron, initially located on one of the three sites, hopping to one or the other of the remaining two sites (assumed to be initially unoccupied) is modified by a contribution which, both in sign and magnitude, is linearly proportional to the applied magnetic field. This effect will be shown to arise from the interference between the amplitude for a direct transition between the initial and final sites, and the amplitude for an indirect, second order transition, involving intermediate occupancy of the third site.

The fundamental equation of motion, upon which subsequent calculation is based, is derived in Section I. This equation represents a generalization (appropriate for a treatment of the Hall effect), of the one-dimensional small-polaron theory of II.

On the basis of this equation, the magnetic field dependent component of the jump probability is calculated for the high temperature site-jump regime ($T > T_t$). Two approaches are used. The first, presented in Section II, is a classical occurrence—probability treatment which applies at temperatures large compared to the Debye temperature. In this treatment, which is the more physically intuitive approach, the lattice displacements are considered as known functions of time. It turns out that the dominant contributions to the Hall effect occur for those three-site configurations which provide a momentary coincidence in the three "local-site" energies (i.e., the electronic energies for local site occupation). The resulting jump rate is expressed as essentially the product of the probability per unit time for the triple coincidence event times the jump probability per coincidence event.

A fully quantum-mechanical calculation of the three-site jump rate is presented in Section III. Here, one considers transitions between states in which the electron is localized at a given site, and the vibrational states described by a set of vibrational quantum numbers, $(\cdots N_k \cdots)$. The computed transition rate displays quantum corrections to the above-mentioned classical result, and agrees with it in the classical, high-temperature limit, thereby verifying the principal physical features of the coincidence approach.

In Section IV, these results are used to calculate the Hall mobility and Hall constant of a two-dimensional lattice in which the three sites are mutually nearest neighbors. The Hall effect is found to be larger than the "normal" Hall effect (Hall constant $\sim -1/nec$); for representative values of the parameters, it exceeds the normal value by a factor ~ 30. The Hall constant appropriate to a square (four-site) lattice arrangement is briefly considered in

Section V. The order of magnitude of this quantity (based on preliminary considerations) is comparable with the "normal" value.

Finally, in Section VI, the Hall effect appropriate to the low temperature ($T < T_t$) polaron-band regime, is discussed briefly. The orders of magnitude of the associated Hall constants are comparable with those found in the above mentioned site-jump regime.

I. THE BASIC EQUATIONS

The mathematical apparatus of the present treatment is based on a suitable generalization of the one-dimensional molecular-crystal model (MCM), previously utilized in I and II. Before developing this generalization, let us briefly recapitulate the essential features of this model.

In the absence of the electron, the system is taken to be a linear chain of N identical diatomic molecules whose orientations and centers of gravity are fixed, but whose internuclear separations are allowed to vary. The "lattice vibrations" of the "host crystal" thus consist of the vibrations of the individual internuclear separations. The "lattice Hamiltonian" is assumed to have the form

$$H_L = \sum_{m=1}^{N} \left(\frac{1}{2M} P_m^2 + \frac{1}{2} M\omega_0^2 x_m^2 \right) + \frac{1}{2} \sum_{m=1}^{N} M\omega_1^2 x_m x_{m+1} \tag{1.1}$$

Here, the x_m are the deviations from equilibrium of the internuclear separations of the individual molecules. The first two terms of (1.1) represent the kinetic and potential energies characteristic of isolated molecules (in the harmonic approximation). The third term provides nearest-neighbor coupling between the x_m ; as is known, such a term (among other possibilities) gives rise to dispersion of vibrational frequencies, which, as discussed at length in II, is a necessary ingredient of small-polaron theory.

The motion of the excess electron (or hole) in the above described molecular-crystal medium is formulated in terms of the tight-binding approximation. For the details of this formulation, the reader is referred to I, Appendix I. Its principal features are as follows. The state of the system is expressed as a linear superposition

$$\psi(\mathbf{r}, x_1 \cdots x_N) = \sum_n a_n(x_1 \cdots x_N)\phi(\mathbf{r} - n\mathbf{a}, x_n) \tag{1.2}$$

of "molecular" electron wave functions $\phi(\mathbf{r} - n\mathbf{a}, x_n)$, each localized about a particular nth molecular site ($\mathbf{a}$ is a unit lattice vector) and depending upon the internuclear coordinate, x_n , of that site. The equation satisfied by a particular $\phi(\mathbf{r} - n\mathbf{a}, x_n)$ is (cf. I, Eq. (I-3))

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r} - n\mathbf{a}, x_n) \right] \phi(\mathbf{r} - n\mathbf{a}, x_n) = E(x_n)\phi(\mathbf{r} - n\mathbf{a}, x_n) \tag{1.3}$$

where $U(\mathbf{r} - n\mathbf{a}, x_n)$ is the contribution of the nth molecule to the effective one-electron potential, and where $E(x_n)$, the eigenvalue of (1.3), generally depends upon x_n.

The coefficients, $a_n(x_1 \cdots x_N)$, of the superposition (1.2) are each functions of all the $x_1, x_2, \cdots x_N$. The equations which they obey are obtained from the time-dependent Schrödinger equation of the system by the standard "projection" procedure. With approximations appropriate to the tight-binding case (smallness of transfer and nonorthogonality integrals), and to the large mass ratio of electrons and nuclei, they take the form[1] (cf. I, Eqs. (I-16), (I-14), (I-15))

$$i\hbar \frac{\partial}{\partial t} a_n(x_1 \cdots x_N) \equiv [H_L + E(x_n) + W_n(x_1 \cdots x_N)]a_n(x_1 \cdots x_N) + \sum_{(\pm)} J(x_n, x_{n\pm1})a_{n\pm1}(x_1 \cdots x_N), \tag{1.4}$$

where

$$W_n(x_1 \cdots x_N) \equiv \int |\phi(\mathbf{r} - n\mathbf{a}, x_n)|^2 \sum_{p \neq n} U(\mathbf{r} - p\mathbf{a}, x_p)\, dV \tag{1.5}$$

$$J(x_n, x_m) \equiv \int \phi(\mathbf{r} - n\mathbf{a}, x_n)U(\mathbf{r} - n\mathbf{a}, x_n)\phi(\mathbf{r} - m\mathbf{a}, x_n)\, dV \tag{1.6}$$

In the detailed calculations of I and II, three simplifications of (1.4) are introduced. These are:

(a) Neglect of the energy terms, $W_n(x_1 \cdots x_N)$.

(b) Neglect of the x_n-dependence of the $J(x_n, x_{n\pm1})$; this simplification means that the $J(x_n, x_{n\pm1})$ are all to be taken equal to a single constant (which, in view of the fact that U is intrinsically negative, shall be written as $-J$, where J is positive).

(c) The x_n-dependence of $E(x_n)$ is taken to be linear, i.e.,

$$E(x_n) = -Ax_n$$

with A a constant (positive or negative); the additive constant $E(0)$ is taken to be zero, again for convenience.[2]

[1] Equation (1.4) differs from (I-16) of I by the inclusion of the intermolecular coupling term present in (1.1)(of this paper).

[2] A discussion of the significance of these simplifications is given in I. Here, it may be remarked that (b) and (c) (which, in essence, constitute linearization approximations, appropriate to small-amplitude vibrational displacements) are also common to other formulations of the problem (cf. J. YAMASHITA AND T. KUROSAWA, *J. Phys. Chem. Solids* **5**, 34 (1958) and G. SEWELL, *Phil. Mag.* **3**, 1361 (1958)). As far as (a) is concerned, its employment leads to the admittedly unphysical feature of a completely local electron-lattice

Incorporating these three assumptions into (1.4), one has

$$i\hbar \frac{\partial}{\partial t} a_n(x_1 \cdots x_N) = (H_L - Ax_n)a_n(x_1 \cdots x_N) - J(a_{n+1} + a_{n-1}) \quad (1.7)$$

which, together with (1.1), constitutes the basic starting point of the small-polaron theory of II.

Before (1.7) can be applied to a study of the Hall effect, it must be generalized in three ways. One of these, the extension to more than one spatial dimension, is trivial; one merely replaces the scalar site index n by a vector index $\mathbf{g} = \sum_i g_i \mathbf{a}_i$, where the $\mathbf{a}_i$ are a basic set of lattice displacement vectors and the g_i are integers. In what follows, the host crystal will be taken to be two-dimensional (two being the minimum number of dimensions required for a theory of the Hall effect) so that $\mathbf{g} = g_1\mathbf{a}_1 + g_2\mathbf{a}_2$. One thus has, in place of (1.7),

$$i\hbar \frac{\partial}{\partial t} a_{\mathbf{g}}(\cdots x_{\mathbf{g}} \cdots) = (H_L - Ax_{\mathbf{g}})a_{\mathbf{g}} - \sum_{\mathbf{h}} a_{\mathbf{g}+\mathbf{h}} J_{\mathbf{h}} \quad (1.8)$$

where the vector $\mathbf{h}$ indexes the nearest neighbors, $\mathbf{g} + \mathbf{h}$, of an arbitrary site, $\mathbf{g}$, and where, in place of (1.1) and (1.6), one has

$$H_L = \sum_{\mathbf{g}} \left(\frac{1}{2M} P_{\mathbf{g}}^2 + \frac{1}{2} M\omega_0^2 x_{\mathbf{g}}^2 + \frac{1}{2} \sum_{\mathbf{h}} M\omega_1^2 x_{\mathbf{g}} x_{\mathbf{g}+\mathbf{h}} \right) \quad (1.9)$$

and

$$J_{\mathbf{h}} = \int \phi^*(\mathbf{r} - \mathbf{g}) U(\mathbf{r} - \mathbf{g}) \phi(\mathbf{r} - \mathbf{g} - \mathbf{h})\, dV, \quad (1.10)$$

with $\phi(\mathbf{r} - \mathbf{g})$ satisfying the equation (obtained from (1.3) by replacing $n\mathbf{a}$ by $\mathbf{g}$, and by setting $x_{\mathbf{g}}$ equal to zero in accordance with assumption (b))

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r} - \mathbf{g}) \right] \phi(\mathbf{r} - \mathbf{g}) = E\phi(\mathbf{r} - \mathbf{g}) \quad (1.11)$$

interaction—the local electronic energy of the nth site, $E(x_n)$, depending on x_n alone—in contrast to the actual situation, in which this local energy would be expected to depend upon the lattice displacements at a number of more or less neighboring sites. Actually, from a purely calculational point of view, assumption (a) is not essential. It turns out that, if a function $W_n(x_1, \cdots x_N)$ (linear in the x_N) were included, the resultant Hamiltonian, when written in terms of normal-mode coordinates (as is done in I, II, and the present work), would have a form very similar to that pertaining to the Hamiltonian of (1.4); the sole difference would be that the constant A (introduced in assumption (c)) would get replaced by a function A_k, of the wave vector, k, of a given lattice-vibration mode. A generalization of this type can be accomodated by the formalism; however, in the opinion of the present authors, its introduction would serve only to render the treatment more complicated, without changing the final results in any significant way.

(E being the eigenvalue appropriate to $x_g = 0$; as in assumption (c), it may be set equal to zero without loss of generality).

The second generalization of (1.7) has to do with the effect of the applied magnetic field. Following a procedure employed by Zilberman (7) one takes account of the magnetic field by introducing new local functions, $\phi_g(\mathbf{r}, x_g)$ which satisfy the equations

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\,\mathrm{grad}_r + \frac{e\mathbf{A}}{c}\right)^2 + U(\mathbf{r} - \mathbf{g}, x_g)\right]\phi_g(\mathbf{r}, x_g) = \phi_g(\mathbf{r}, x_g) \tag{1.12}$$

where, for convenience, the "symmetrical" gauge

$$\mathbf{A} = \tfrac{1}{2}[\mathbf{H} \times \mathbf{r}] \tag{1.13}$$

is employed. The analysis of Appendix I of I is carried out just as before (the sole exception being that the original functions $\phi(\mathbf{r} - \mathbf{g}, x_g)$, which apply for the case of zero magnetic field, are replaced by their field-dependent counterparts, $\phi_g(\mathbf{r}, x_g)$). Upon incorporating assumptions (a), (b), and (c) into the resulting equations, one reobtains (1.8) and (1.9); in place of (1.10), however, one has

$$J_{g,g+h} = \int \phi_g^*(\mathbf{r}) U(\mathbf{r} - \mathbf{g}) \phi_{g+h}\, dV \tag{1.14}$$

where $\phi_g(\mathbf{r})$ represents the function $\phi_g(\mathbf{r}, x_g)$ evaluated at $x_g = 0$ (in accordance with assumption (b)); it satisfies the equation

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\,\mathrm{grad}_r + \frac{e\mathbf{A}}{c}\right)^2 + U(\mathbf{r} - \mathbf{g})\right]\phi_g(\mathbf{r}) = E\phi_g(\mathbf{r}) \tag{1.15}$$

(obtained from (1.12) by taking x_g equal to zero).

A more useful form of this equation is achieved via the introduction of the gauge transformation

$$\phi_g(\mathbf{r}) = \chi_g(\mathbf{r}) e^{-ie[\mathbf{H}\times\mathbf{g}]\cdot\mathbf{r}/2\hbar c} \tag{1.16}$$

One then has

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\,\mathrm{grad}_r + \frac{e}{2c}\mathbf{H} \times (\mathbf{r} - \mathbf{g})\right)^2 + U(\mathbf{r} - \mathbf{g})\right]\chi_g(\mathbf{r}) = E\chi_g(\mathbf{r}) \tag{1.17}$$

the solutions of which are obviously of the form

$$\chi_g(\mathbf{r}) = \chi(\mathbf{r} - \mathbf{g}) \tag{1.18}$$

At this point it is desirable to introduce into the two-dimensional MCM an additional simplifying feature. Specifically, it will now be assumed that both the external magnetic field and the vibration axes of the diatomic molecules are

perpendicular to the plane of the two-dimensional crystal. It then follows, immediately, that $U(\mathbf{r} - \mathbf{g})$ possesses cylindrical symmetry about an axis (say the z-axis) perpendicular to this plane, so that the corresponding component of angular momentum (pertaining to a local wave-function), l_z, is a constant of the motion. It then follows that, if the local solution $\chi(\mathbf{r} - \mathbf{g})$ be assumed nondegenerate, it is automatically a Σ-state ($l_z = 0$). With H also in the z-direction, the magnetic-field perturbation in (1.14) takes the form

$$H_{\text{MAGN}} = \frac{eH}{2mc} l_z + O(H^2) \tag{1.19}$$

from which one immediately sees that, by virtue of the fact that $l_z\chi(\mathbf{r} - \mathbf{g}) = 0$, the $\chi(\mathbf{r} - \mathbf{g})$ differ from the zero-field local functions $\phi(\mathbf{r} - \mathbf{g})$ by, at most, terms in H^2 (or higher powers of H). Such corrections are irrelevant for the theory of the present paper, which is limited to effects linear in H, and will henceforth be ignored. One thus has, to the required degree of accuracy

$$\phi_{\mathbf{g}}(\mathbf{r}) = e^{-ie[\mathbf{H}\times\mathbf{g}]\cdot\mathbf{r}/2\hbar c}\phi(\mathbf{r} - \mathbf{g}) \tag{1.20}$$

It may here be mentioned, in passing, that the above-described prescriptions concerning the directions of magnetic field and molecular axes are introduced solely for the purpose of simulating the situation to be expected in a tight-binding treatment of a real ionic crystal, in which the wave functions of non-degenerate local-site levels are s-functions, and hence unaffected, in first order, by the presence of the magnetic field.

The dependence of $J_{\mathbf{g},\mathbf{g}+\mathbf{h}}$ on the magnetic field may now be determined. Substituting (1.20) into (1.14), and noting that the zero-field functions are real, one has

$$J_{\mathbf{g},\mathbf{g}+\mathbf{h}} = \int \phi(\mathbf{r} - \mathbf{g})\phi(\mathbf{r} - \mathbf{g} - \mathbf{h})U(\mathbf{r} - \mathbf{g})e^{-ie[\mathbf{H}\times\mathbf{h}]\cdot\mathbf{r}/2\hbar c}\, dV$$

Transforming the integration variable to $\boldsymbol{\varrho} = \mathbf{r} - \mathbf{g} - \mathbf{h}/2$ and expanding the exponent about $\boldsymbol{\varrho} = 0$ to terms linear in H, one has

$$J_{\mathbf{g},\mathbf{g}+\mathbf{h}} = e^{-ie\mathbf{H}\cdot[\mathbf{h}\times\mathbf{g}]/2\hbar c} \int \phi\left(\boldsymbol{\varrho} + \frac{\mathbf{h}}{2}\right)\phi\left(\boldsymbol{\varrho} - \frac{\mathbf{h}}{2}\right)U\left(\boldsymbol{\varrho} + \frac{\mathbf{h}}{2}\right) \times\left[1 - \frac{ie\mathbf{H}\cdot[\mathbf{h}\times\boldsymbol{\varrho}]}{2\hbar c}\right] dV \tag{1.21}$$

It may now be noted that, by virtue of the cylindrical symmetry of $\phi(\mathbf{r})$ and $U(\mathbf{r})$ (about the z-axis, perpendicular to the plane of the two-dimensional crystal), the component of the vector

$$\int \phi\left(\varrho + \frac{\mathbf{h}}{2}\right) U\left(\varrho + \frac{\mathbf{h}}{2}\right) \phi\left(\varrho - \frac{\mathbf{h}}{2}\right) \varrho \, dV$$

which lies in the plane of the crystal must be directed along the intersite axis,[3] $\mathbf{h}$, and hence cannot contribute to (1.21). One thus has

$$J_{\mathbf{g},\mathbf{g}+\mathbf{h}} = -Je^{i\alpha_{\mathbf{g},\mathbf{g}+\mathbf{h}}} \tag{1.22}$$

where (by virtue of the assumed symmetry properties of $\phi(\mathbf{r})$ and $U(\mathbf{r})$)

$$J = \int \phi(\mathbf{r}) U(\mathbf{r}) \phi(\mathbf{r} - \mathbf{h}) \, dV \tag{1.23}$$

is a positive constant for all nearest neighbors, and where

$$\alpha_{\mathbf{g},\mathbf{g}+\mathbf{h}} = -\frac{e}{2\hbar c} \mathbf{H}\cdot[\mathbf{h} \times \mathbf{g}] = -\frac{e}{\hbar c} \mathbf{H}\cdot\mathbf{A}_{\mathbf{g},\mathbf{g}+\mathbf{h}} \tag{1.24}$$

with

$$\mathbf{A}_{\mathbf{g},\mathbf{g}+\mathbf{h}} \equiv \tfrac{1}{2}[\mathbf{h} \times \mathbf{g}] = \tfrac{1}{2}[(\mathbf{g} + \mathbf{h}) \times \mathbf{g}];$$

$\mathbf{A}_{\mathbf{g},\mathbf{g}+\mathbf{h}}$ is clearly the area of the triangle, two of whose vertices coincide with the centroids of sites $\mathbf{g}$ and $\mathbf{g} + \mathbf{h}$, the third vertex being the origin.

It will now be noted that, although the "magnetic phase factors," $\alpha_{\mathbf{g},\mathbf{g}'}$ are not uniquely defined (by virtue of the arbitrariness of the origin), the sum of the phase factors linking any set of three (or more) sites in a closed cycle, is determined. Specifically, one has

$$\alpha_{\mathbf{g}_2\mathbf{g}_1} + \alpha_{\mathbf{g}_3\mathbf{g}_2} + \alpha_{\mathbf{g}_1\mathbf{g}_3} = -\frac{e}{\hbar c} \mathbf{H}\cdot\mathbf{A}_{321} \tag{1.25}$$

where

$$\mathbf{A}_{321} \equiv \tfrac{1}{2}[(\mathbf{g}_1 \times \mathbf{g}_2) + (\mathbf{g}_2 \times \mathbf{g}_3) + (\mathbf{g}_3 \times \mathbf{g}_1)] \tag{1.26}$$

is the area of the triangle subtended by the centroids of the three sites (described in the sense $1 \to 2 \to 3 \to 1$); this fact is easily established by introducing the intersite displacement vectors

$$\mathbf{h}_{ji} \equiv \mathbf{g}_j - \mathbf{g}_i \tag{1.27}$$

in terms of which (1.26) takes the more familiar form

$$\mathbf{A}_{321} = \tfrac{1}{2}(\mathbf{h}_{21} \times \mathbf{h}_{31}) \tag{1.28}$$

It should here be remarked that, although the basic result (1.25), was es-

[3] If, say, the x-axis be chosen parallel to $\mathbf{h}$, the y-component of the integrand is of the form $f(x, y)y$, where $f(x, y) = f(x, -y)$. The integral of the y-component is then obviously zero, q.e.d.

tablished through the use of a particular gauge, namely, the symmetrical one (given by (1.13)), one can obtain this result by an alternate method in which the gauge is arbitrary. This treatment is given in Appendix E of the present paper. Specifically, it is demonstrated in this appendix that although the phases of the various $J_{\mathbf{g},\mathbf{g}+\mathbf{h}}$, (1.22) are gauge dependent, the result of physical interest, (1.25), is properly gauge invariant.

The third generalization of (1.7) involves taking into account the existence of a spatially constant dc electric field, $\mathbf{F}$. The procedure is straightforward; one merely applies the procedure of Appendix I of I to the previous Hamiltonian, augmented by the standard electric-field term

$$H^{(F)} = e\mathbf{F}\cdot\mathbf{r}$$

In this process, $H^{(F)}$ automatically separates into two parts. One of these, namely $e\mathbf{F}\cdot\mathbf{g}$ appears on an additive modification of the local-site energy term, $-Ax_{\mathbf{g}}$. The other, $e\mathbf{F}\cdot(\mathbf{r}-\mathbf{g})$ gets lumped together with the local-site Hamiltonian[4] (given by the l.h.s. of (1.15)).

The results of the above-discussed three generalizations may be summarized by writing down the equation obeyed by the coefficients $a_{\mathbf{g}}(\cdots x_{\mathbf{g}}\cdots)$ of the total wave function. It is (cf. (1.8), (1.22), and preceding paragraph)

$$i\hbar\frac{\partial a_{\mathbf{g}}}{\partial t} = (H_{\mathrm{L}} - Ax_{\mathbf{g}} + e\mathbf{F}\cdot\mathbf{g})a_{\mathbf{g}} - J\sum_{\mathbf{h}} a_{\mathbf{g}+\mathbf{h}}\, e^{i\alpha_{\mathbf{g},\mathbf{g}+\mathbf{h}}} \tag{1.29}$$

where H_{L} is given by (1.9) and the $\alpha_{\mathbf{g},\mathbf{g}'}$ may be considered as defined by the totality of independent cyclical relations of the form of (1.26).

Equation (1.29) constitutes the basis of the treatment of this paper; its application to various cases of interest is presented in subsequent sections.

II. CLASSICAL OCCURRENCE-PROBABILITY APPROACH

To begin, some general geometric features will first be noted: Except for a square lattice arrangement to be considered in Section V, the elementary transition probabilities are to be calculated for the three mutually near neighbor site arrangement shown in Fig. 1. The relative site vectors joining the site centroids form an equilateral triangle as shown. The origin is taken at the final site, $\mathbf{g}_3$. Evidently, $\mathbf{h}_{31} = \mathbf{h}_{32} + \mathbf{h}_{21}$.

Next, before proceeding with the calculation, it may be helpful to point out again that the present occurrence-probability approach is the more physically intuitive one, the principal result of this section being given by (2.19). A more complete quantum-mechanical approach, presented in Section III, is more mathematical in nature. The result of this approach is identical with (2.19)

[4] In principle, the local functions, and hence the parameters, J, acquire a small electric-field dependence; this higher order correction will be ignored in the present work.

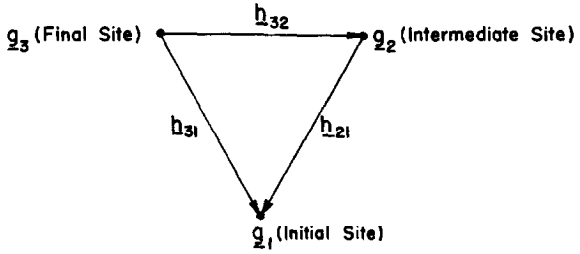

FIG. 1. Elementary three-site configuration

in the classical limit ($T \gg \Theta$), but, in addition, exhibits certain quantum-mechanical corrections of interest in the nonclassical domain.

The starting point of the classical occurrence-probability approach is the set of equations:

$$i\hbar \frac{\partial a_{\mathbf{g}}}{\partial t} = [-Ax_{\mathbf{g}} + e(\mathbf{F}\cdot\mathbf{g})]a_{\mathbf{g}} - J \sum_{\underline{\mathbf{h}}} e^{+i\alpha_{\mathbf{g},\mathbf{g}+\mathbf{h}}}\, a_{\mathbf{g}+\mathbf{h}} \tag{2.1}$$

which is obtained from (1.29) by omitting the vibrational Hamiltonian, H_{L}. This omission is based on the fact that in this method the lattice vibratory motion is not treated quantum mechanically, but, rather, the $x_{\mathbf{g}}(t)$ are regarded as given functions of time. Specifically, the $x_{\mathbf{g}}(t)$ are determined by the Hamiltonian $H^{(\mathrm{initial})} = H_{\mathrm{L}} - Ax_{(\mathrm{initial})}$, i.e., the sum of the purely vibrational Hamiltonian and interaction Hamiltonian *before* making the transition to a neighboring site. The question of whether $H^{(\mathrm{initial})}$ or $H^{(\mathrm{final})}$ should be used in determining the $x_{\mathbf{g}}(t)$ causes no difficulty within the domain of validity of this approach, as shown in I (Appendix II).[5]

In proceeding with the analysis of (2.1), it will be convenient to drop the electric-field term; it can be reinserted, when desired, by adding it to the electronic energy, according to the prescription

$$-Ax_{\mathbf{g}} \rightarrow -Ax_{\mathbf{g}} + e\mathbf{F}\cdot\mathbf{g}$$

At some initial time, $t = -T/2$, the electron is taken to be localized at site $\mathbf{g}_1$. In the absence of overlap ($J = 0$) the electron would remain on this site, its occupation amplitude being given by[6]

$$a_1^{(0)} = e^{i\phi_i(t)} \tag{2.2}$$

[5] Basically, the reason is that, as will be seen below, the effective time-duration of energy coincidence events, during which site-jumps occur with appreciable net probability, turns out to be small enough so that vibratory velocities do not have time to change, and are in fact, considered as constants for a given coincidence event. The change in the effective vibrational Hamiltonian (due to change of electronic state) is therefore unimportant.

[6] Herein, and in what follows, site-subscripts of the form $\mathbf{g}_1$, $\mathbf{g}_2$, etc., will be replaced by subscripts, 1, 2, etc. Thus, $a_{g_i} \rightarrow a_i$, $x_{g_i} \rightarrow x_i$, $\alpha_{g_j,g_i} \rightarrow \alpha_{ji}$.

where

$$\phi_i(t) \equiv \frac{A}{\hbar} \int_0^t x_i(t')\, dt' \tag{2.3}$$

for a given site, $\mathbf{g}_1$. (The phases are arbitrarily taken to be zero at $t = 0$).

In the presence of nonvanishing (but small) J, one obtains the standard perturbation development for the occupation amplitudes of the other sites; in particular, for the three-site configuration of Fig. 2, one has

$$a_3^{(1)}(t) = -\frac{J}{i\hbar} e^{i\alpha_{31}} e^{i\phi_3(t)} \int_{-T/2}^{t} e^{i[\phi_1(t')-\phi_3(t')]}\, dt' \tag{2.4}$$

$$\begin{aligned} a_3^{(2)}(t) &= -\frac{J}{i\hbar} e^{i\alpha_{32}} e^{i\phi_3(t)} \int_{-T/2}^{t} a_2^{(1)}(t') e^{-i\phi_3(t')}\, dt' \\ &= -\frac{J^2}{\hbar^2} e^{i[\alpha_{32}+\alpha_{21}]} e^{i\phi_3(t)} \int_{-T/2}^{t} e^{i[\phi_2(t')-\phi_3(t')]}\, dt' \int_{-T/2}^{t'} e^{i[\phi_1(t'')-\phi_2(t'')]}\, dt'' \end{aligned} \tag{2.5}$$

The occupation probability, $W_3(t)$, of site $\mathbf{g}_3$ at time t is obtained by taking the absolute square of $a_3^{(1)} + a_3^{(2)}$. For the purposes of the present paper, however, only that component, $W_3^{(H)}(t)$, which is *linearly* proportional to the magnetic field, is of interest; it alone can contribute to a nonvanishing Hall effect. To obtain $W_3^{(H)}(t)$, one expands the $e^{i\alpha_{ji}}$ as power series in the α_{ji}, and discards all terms except those linear in the α_{ji}. One then has (upon setting t equal to $T/2$ so that the total time interval is T)

$$\begin{aligned} W_3^{(H)}\left(\frac{T}{2}\right) = \frac{\alpha J^3}{\hbar^3} \Bigg[\Bigg(\int_{-T/2}^{T/2} e^{-i[\phi_1(t')-\phi_3(t')]}\, dt' \Bigg) \Bigg(\int_{-T/2}^{T/2} e^{i[\phi_2(t')-\phi_3(t')]}\, dt' \\ \times \int_{-T/2}^{t'} e^{i[\phi_1(t'')-\phi_2(t'')]}\, dt'' \Bigg) + \text{c.c.} \Bigg] \end{aligned} \tag{2.6}$$

where[7]

$$\alpha \equiv \alpha_{21} + \alpha_{32} + \alpha_{13} \equiv \frac{e\mathbf{H}\cdot\mathbf{A}_{321}}{\hbar c} \tag{2.7}$$

The principal task is that of taking a thermal average of (2.6) over all vibrational "trajectories" (i.e., over all dynamically permissible time variations of the x_i). Before getting into detailed calculations, one may conveniently take advantage of a general property of the vibrational trajectories to effect a considerable simplification. The property in question is that of time-reversal. Namely, for every trajectory, characterized by three functions $x_i(t)$, $(i = 1, 2, 3)$, there exists a corresponding "time-reversed" trajectory, characterized by the three functions.

[7] In obtaining (2.7), use is made of the formula, $\alpha_{31} = -\alpha_{13}$.

$$x_i^{(\mathrm{tr})}(t) = x_i(-t), \qquad (i = 1, 2, 3) \quad (2.8)$$

(where the superscript "(tr)" denotes the time-reversed trajectory). Inserting (2.8) into (2.3), one has

$$\phi_i^{(\mathrm{tr})}(t) = -\phi_i(-t) \quad (2.9)$$

The simplification consists in "time-reversal symmetrization" of[8] (2.6). Specifically, one adds to the r.h.s. of (2.6) the expression gotten by replacing each $\phi_i(t)$ by $\phi_i^{(\mathrm{tr})}(t)$, and divides by 2. Then, upon employing (2.9), one has

$$W_3^{(H)}\left(\frac{T}{2}\right) = \frac{\alpha J^3}{2\hbar^3}\left[\left(\int_{-T/2}^{+T/2} e^{-i[\phi_1(t')-\phi_3(t')]}\,dt'\right)\left(\int_{-T/2}^{+T/2} e^{i[\phi_2(t')-\phi_3(t')]}\,dt'\right.\right.$$
$$\times \left.\int_{-T/2}^{t'} e^{i[\phi_1(t'')-\phi_2(t'')]}\,dt''\right) + \left(\int_{-T/2}^{+T/2} e^{i[\phi_1(-t')-\phi_3(-t')]}\,dt'\right) \quad (2.10)$$
$$\times \left.\left(\int_{-T/2}^{+T/2} e^{-i[\phi_2(-t')-\phi_3(-t')]}\,dt' \int_{-T/2}^{t'} e^{-i[\phi_1(-t'')-\phi_2(-t'')}\,dt''\right)\right] + \text{c.c.}$$

where the subscript, "trs", is used to denote the result of the above described time-reversal symmetrization procedure.

It will now be noted that, upon introducing a transformation in which the signs of the integration variables, t' and t'', in the second term of (2.10) are changed, one obtains an expression in which the various integrands are all complex conjugates of corresponding integrands in the first term of (2.10); moreover, the integration intervals, $-T/2 \to T/2$ remain unchanged, whereas the interval, $-T/2 \to t'$ (for the t'' integration) goes over into $t' \to T/2$. The first term can then be conveniently added to the complex conjugate of the second to yield:

$$W_3^{(H)}\left(\frac{T}{2}\right) = \frac{\alpha J^3}{2\hbar^3}\left[\left(\int_{-T/2}^{-T/2} dt e^{i[\phi_1(t)-\phi_2(t)]}\right)\left(\int_{-T/2}^{T/2} dt' e^{i[\phi_2(t')-\phi_3(t')]}\right)\right.$$
$$\left.\times\left(\int_{-T/2}^{T/2} dt'' e^{i[\phi_3(t'')-\phi_1(t'')]}\right)\right] + \text{c.c.} \quad (2.11)$$

The present classical-coincidence treatment, as discussed in (II, Appendix II), is based on the fact that with the conditions[9]

$$\kappa T \gg \hbar\omega_0\,, \qquad \frac{A^2}{M\omega_0^2} \gg \hbar\omega_0\,, \quad (2.12)$$

[8] The simplification is made possible by the fact that the classical occurrence-probability for a given vibrational trajectory is identical (as it must be) with the probability corresponding to the time-reversed path. This will be later verified in Appendix A. (See footnote 53.)

[9] The first inequality simply represents the regime in which the lattice motion can be treated classically; the second insures that the phase undergo many oscillations between successive coincidence events.

the oscillations of the exponential integrands interfere destructively, except at points of coincidence of the *electronic* energies (at which $x_i = x_{i+1}$). At such points, the phases of the above exponentials are momentarily stationary.

Let us assume that, within the interval $-T/2 \leqq t \leqq T/2$, one such point[10] of stationary phase occurs between sites 1 and 3 (see Fig. 1) at some arbitrary instant which is conveniently taken to be $t = t_{31} = 0$, without loss of generality. In addition, assume that a coincidence of sites 1 and 2 takes place at $t = t_{21}$ within the interval $-T/2$ to $T/2$.

Further, the explicit time dependence of the relative displacements,

$$(x_i(t) - x_{i+1}(t)),$$

required for the evaluation of the integrals appearing in (2.11) is assumed to be linear.[11] Denoting the relative velocity between sites i and j by

$$v_{ij} = v_i - v_j , \tag{2.13}$$

one may write[12]

$$\begin{aligned} x_1 - x_2 &= v_{12}(t - t_{12}) \\ x_3 - x_1 &= v_{31}t \end{aligned} \tag{2.14}$$

Substituting (2.14) and (2.3) into (2.11), one obtains

$$\begin{aligned} W_3^{(H)}\left(\frac{T}{2}\right) = \frac{\alpha J^3}{2\hbar^3} \exp\left(\frac{iA}{2\hbar}\, v_{12}\frac{v_{31}}{v_{23}}\, t_{12}^2\right) \int_{-T/2}^{T/2} dt \exp\left[\frac{iA}{2\hbar}\, v_{12}(t - t_{12})^2\right] \\ \times \int_{-T/2}^{T/2} dt' \exp\left[\frac{iA}{2\hbar}\, v_{23}(t' - t_{23})^2\right] \times \int_{-T/2}^{T/2} dt'' \exp\left(\frac{iA}{2\hbar}\, v_{31}\, t''^2\right) + \text{c.c.} \end{aligned} \tag{2.15}$$

[10] For this assumption to be valid, it is necessary that T be small compared to the time Δt_c between successive coincidence of sites 1 and 3. According to II (bottom of p. 385), a lower estimate of Δt_c is

$$\Delta t_c \sim \frac{1}{\omega_0}\left(\frac{2\kappa T}{A^2/M\omega_0^2}\right)^{1/2}$$

On the other hand, as shown immediately below, it is necessary that T be larger than a characteristic interval $\sim(\hbar/Av_{\text{th}})^{1/2}$. These conditions are simultaneously realizable when

$$\kappa T \gg \left[\frac{A^2}{M\omega_0^2}\,(\hbar\omega_0)^2\right]^{1/3}$$

[11] As discussed in (II, Appendix II), the validity of the linear approximation and the neglect of acceleration terms requires that the relative acceleration times the effective time duration of the coincidence event, be much less than the mean (thermal) velocity. This requirement places a lower limit on the temperature as given by the inequality

$$T \gg \left[\frac{A^2}{M\omega_0^2}\,(\hbar\omega_0)^2\right]^{1/3},$$

(which was already cited in footnote 7).

[12] Note that $(x_2 - x_3)$ is not independent; rather $v_{23} = -(v_{12} + v_{31})$ and $t_{23} = -(v_{12}/v_{23})t_{12}$.

It is first to be noted that the dominant contributions to the individual integrals come from intervals $\sim(\hbar/Av_{\text{th}})^{1/2}(\ll T/2)$ about their respective points of stationary phase. Now, as shown immediately below ((2.17) et seq.), the principal contributions to the net (magnetic) field-dependent jump-rate (after averaging over all configurations) occur for values of $t_{12} \leqq (\hbar/Av_{\text{th}})^{1/2}$. (Barring exceptionally large values of the ratio v_{12}/v_{23}, one then also has that $t_{23} \lesssim (\hbar/Av_{\text{th}})^{1/2}$.) From the inequality $T \gg (\hbar/Av_{\text{th}})^{1/2}$ given in footnote 14, it then follows that all three coincidence points are contained within the interval $-T/2 \leqq t \leqq T/2$.

The individual limits of integration can then be extended from $-\infty$ to ∞, with the result:

$$W_3^{(H)} = \frac{\alpha J^3}{2\hbar^3} \exp\left(\frac{iA}{2\hbar} v_{12} \frac{v_{31}}{v_{23}} t_{12}^2\right)\left[\left(\frac{2\pi\hbar}{A \mid v_{12} \mid}\right)^{1/2} e^{\pm i\pi/4}\right]\left(\frac{2\pi\hbar}{A \mid v_{23} \mid}\right)^{1/2} e^{\pm i\pi/4}\Bigg] \times \left[\left(\frac{2\pi\hbar}{A \mid v_{31} \mid}\right)^{1/2} e^{\pm i\pi/4}\right] + \text{c.c.} \tag{2.16}$$

where $(\pm)$ in each bracket is taken according to whether $v_{ij}(\gtrless)0$.

The result represents the transition probability for the previously described triple coincidence event. In order to obtain the jump rate, (2.16) must be multiplied by the mean probability for a coincidence of x_1 and x_3 at time $t_{31} = 0$ with relative velocity between v_{31} and $v_{31} + dv_{31}$, and a coincidence of x_1 and x_2 occurring between times t_{12} and $t_{12} + dt_{12}$ with relative velocity between v_{12} and $v_{12} + dv_{12}$. This occurrence-probability is denoted by $P_T{}^C(v_{31}, v_{12}, t_{12})\, dv_{31}\, dv_{12}\, dt_{12}$. Subsequent integrations over v_{31}, v_{12}, and t_{12} then give the desired jump rate, i.e.,

$$w_3^{(H)} = \iiint dv_{31}\, dv_{12}\, dt_{12}\, W_3^{(H)} P_T{}^C(v_{31}, v_{12}, t_{12}) \tag{2.17}$$

It is expedient to consider first the integration over t_{12}. From (2.16), one notes that $W_3^{(H)}$, regarded as a function of t_{12}, has a characteristic width[13] $\Delta t_{12} \sim (\hbar/Av_{\text{th}})^{1/2}$ about its point of stationary phase at $t_{12} = 0$. Beyond this interval, its increasingly rapid oscillations give rise to destructive interference. It then follows that, if $P_T{}^C$ is relatively constant over an interval $\sim\Delta t_{12}$ (about $t_{12} = 0$), one may replace $P_T{}^C$ by its value at $t_{12} = 0$. The detailed calculation of $P_T{}^C$ is carried out in Appendix A; relatively simple expressions are obtained for the case of small dispersion of the vibrational frequencies,

$$\frac{\omega_k - \omega_0}{\omega_0} \ll 1 \qquad \text{("narrow-band" case).}$$

[13] As a result, it may be noted, as a matter of interest, that the only jumps to the intermediate site $(1 \to 2)$ which are "coherent" (in the sense that they contribute significantly to (2.17)) are those which occur within the interval $\sim(\hbar/Av_{\text{th}})^{1/2}$ over which the direct jump $(1 \to 3)$ occurs.

This case is considered for small t_{12} (i.e., $\omega_0 t_{12} \ll 1$). The corresponding occurrence-probability is given by (A.9) and is quoted here for convenience:

$$P_T{}^C(v_{31}, v_{12}, t_{12}) \sim \exp\left(-\frac{\epsilon_2}{\kappa T} - \frac{M v_{31}^2}{4\kappa T} + \mu + \nu\right) \tag{A.9}$$

Here,

$$\mu = -\frac{\epsilon_2}{3\kappa T} - \frac{4}{3}\frac{M v_{31}^2}{\kappa T}(\omega_0 t_{12})^2 + \frac{A v_{31} t_{12}}{6\kappa T},$$

$$\nu = C_1 \frac{\epsilon_2}{\kappa T}(\omega_0 t_{12})^2 - C_2 \frac{\frac{1}{2}(A v_{31} t_{12}) - A v_{12} t_{12}}{\kappa T} - \frac{M v_{12} v_{31}}{3\kappa T} - \frac{M v_{31}^2}{12\kappa T} - \frac{M v_{12}^2}{3\kappa T}$$

where: $\epsilon_2 = A^2/4M\omega_0{}^2$ is the activation energy corresponding to an energetic coincidence of two sites,[14] C_1 and C_2 are numerical coefficients of the order of unity, and terms of the order $(\omega_0 t_{12})^2$ have been neglected in comparison with unity.

It is noted that (A.9) contains two characteristic times over which it shows appreciable variation. The first such time is simply $\sim 1/\omega_0$. The second time, contained in the last term of μ and the second term of ν, is smaller, and is

$$\sim (1/\omega_0)[\kappa T/(A^2/M\omega_0{}^2)]^{1/2}$$

(replacing, as usual, the relative velocities by their thermal average values $(2\kappa T/M)^{1/2}$). However, it is readily verified, using inequalities (2.12) and the one given in footnote (15), that

$$\Delta t_{12} \ll \frac{1}{\omega_0}, \qquad \frac{1}{\omega_0}\left(\frac{2\kappa T}{A^2/M\omega_0{}^2}\right)^{1/2}$$

It then follows, as claimed above, that $P_T{}^C$ does not vary appreciably for $t_{12} \lesssim \Delta t_{12}$, and may therefore be approximated by its value at $t_{12} = 0$. This result, given by (A.8) of Appendix A, is quoted here for convenience:

$$P_T{}^C(v_{31}, v_{12}, 0) = \frac{A^2}{2\pi}\frac{M}{2\pi\sqrt{3}\kappa T}\frac{1}{3\sqrt{3}\kappa T \epsilon_3} \cdot \exp\left[-\frac{\epsilon_3}{\kappa T} - \frac{M}{6\kappa T}(v_{12}^2 + v_{23}^2 + v_{31}^2)\right] |v_{12}| \, |v_{31}| \tag{A.8}$$

[14] It is shown in appendix B that ϵ_2 and ϵ_3 can be interpreted physically as the minimum energies (in excess of the binding energy, $-E_b$) of the two and three site coincidence configuration, respectively. It is also noted that $\epsilon_3 < 2\epsilon_2$, the latter energy corresponding to two pairs of uncorrelated coincidences. A physical argument for this inequality has been given by C. Herring (private conversation), who, in addition, arrived at the conclusion that in many cases, the Hall mobility may exceed the drift mobility. This conclusion is in accord with the results of the present paper.

where $\epsilon_3 = \frac{4}{3}\epsilon_2$ is the activation energy corresponding to an energetic coincidence of three sites.[14]

Carrying out the integration over t_{12}, (2.17) then becomes:

$$w_3^{(H)} = \alpha \frac{J^3}{\hbar^3} \left(\frac{2\pi\hbar}{A} \left| \frac{v_{23}}{v_{12} v_{31}} \right| \right)^{1/2} \left(\frac{2\pi\hbar}{A \mid v_{12} \mid} \right)^{1/2} \left(\frac{2\pi\hbar}{A \mid v_{23} \mid} \right)^{1/2} \left(\frac{2\pi\hbar}{A \mid v_{31} \mid} \right)^{1/2}$$
$$\times \frac{A^2}{2\pi} \frac{M}{2\pi\sqrt{3}\kappa T} \frac{1}{3\sqrt{3}\kappa T \epsilon_3} \tag{2.18}$$
$$\times \exp\left[-\frac{\epsilon_3}{\kappa T} - \frac{M}{6\kappa T} (v_{12}^2 + v_{23}^2 + v_{31}^2) \right] \mid v_{12} \mid \mid v_{31} \mid$$

Finally, the integrations over the relative velocities are carried out by introducing the change of variables

$$v_{31} = \frac{1}{\sqrt{2}} (v + V)$$

$$v_{12} = \frac{1}{\sqrt{2}} (v - V)$$

The final result reads:

$$w_3^{(H)} = \left(\frac{e\mathbf{H} \cdot \mathbf{A}_{321}}{\hbar c} \right) \frac{J^3}{\hbar} \frac{2\pi}{3\sqrt{3}\kappa T \epsilon_3} e^{-\epsilon_3/\kappa T} \tag{2.19}$$

For later calculational purposes, it is useful to write down the modification of the two- and three-site transition rates due to an applied infinitesimal electric field. The inclusion of the electric field in the Hamiltonian was discussed in the text preceding (1.29). In the present calculation, it enters into the occurrence-probability in that one looks for coincidences of the electronic plus electric field energies of neighboring sites, i.e.,

$$\delta(x_i - x_j) \to \delta\left(x_i - x_j - \frac{e\mathbf{F} \cdot \mathbf{h}_{ij}}{A} \right)$$

The modified transition rates are:

$$\begin{aligned} w_2(1 \to 3) &= e^{e\mathbf{F} \cdot \mathbf{h}_{31}/2\kappa T} w_2 , \\ w_3^{(H)}(1 \to 3) &= e^{e\mathbf{F} \cdot (\mathbf{h}_{31} + \mathbf{h}_{21})/3\kappa T} w_3^{(H)} \end{aligned} \tag{2.20}$$

where $w_3^{(H)}$ is given by (2.19) and w_2 is the elementary two-site jump rate derived in (II, (79)):

$$w_2 = \frac{J^2}{\hbar} \left(\frac{\pi}{4\kappa T \epsilon_2} \right)^{1/2} e^{-\epsilon_2/\kappa T} \tag{2.21}$$

It is noted that the effect of the electric field is simply to enhance the jump probability in the direction[15] of the applied field; in the three-site case, the enhancement is a function of the electric-field-induced charge in the energies of both the intermediate and final sites (relative to that of the initial site).

Before proceeding to the quantum mechanical treatment, an alternate classical-dynamical approach[16] will be briefly outlined. This alternate method simply serves as a check on the final result.

The method essentially starts with (2.11). Instead of using the linear approximations (2.14), however, one expands the displacements in normal mode coordinates of the host crystal:[17]

$$x_i(t) = \left(\frac{2}{N}\right)^{1/2} \sum_{\mathbf{k}} q_{\mathbf{k}}(t) \sin (\mathbf{k}\cdot\mathbf{g}_i + \tfrac{1}{4}\pi)$$

Next, the harmonic time dependence of the normal mode coordinates is explicitly taken into account by writing[18]:

$$q_{\mathbf{k}}(t) = Q_{\mathbf{k}} \cos (\omega_{\mathbf{k}} t + \delta_{\mathbf{k}}) + q_{\mathbf{k}}^{(1)}$$

where the $\delta_{\mathbf{k}}$ are arbitrary phase factors.

The remainder of the calculation consists of taking a thermal average over the amplitudes, $Q_{\mathbf{k}}$, and an unweighted average over the phase factors, $\delta_{\mathbf{k}}$, very much like the classical occurrence-probability calculation given in Appendix A of the present paper. The final result of this procedure agrees with (2.20).

III. QUANTUM-MECHANICAL FORMULATION

In the previous sections, the lattice vibratory motion was treated classically in that the $x_{\mathbf{g}}(t)$ were assumed to follow classical, dynamical trajectories as determined by $(H_{\mathrm{L}} + H_{\mathrm{int}})$. In the present section, however, a full quantum mechanical treatment of the problem will be given.

Starting with the augmented Hamiltonian[19] (1.29), the zeroth order eigenstates and matrix elements of the perturbation will first briefly be summarized. To begin, it is desirable to express the vibrational coordinates $x_{\mathbf{g}}$ in terms of the

[15] In the case of an electron, with relative vectors defined as directed from the final to the initial site, the enhancement is opposite the direction of **F**, as it should be.

[16] For calculational details, the reader is referred to III, appendix A.

[17] We employ the transformation coefficients of Pekar and Buimistrov, *Soviet Phys.—Tech. Phys.* **2**, 2478 (1957).

[18] The normal mode displacements, $q_k^{(1)}$, defined in (3.6), arise from the fact that the electron is initially self trapped on site $\mathbf{g}_1$.

[19] The electric field energy term is neglected in order to avoid undue calculational complexity. The field is taken to be infinitesimal, and slightly enhances the probability of a jump in the field direction. It is felt that this feature is adequately treated in the previously discussed classical approximation.

normal mode coordinates of the host crystal, $q_{\mathbf{k}}$. This is done by means of the two-dimensional analog of II, (II-12), namely:

$$x_g = \left(\frac{2}{N}\right)^{1/2} \sum_{\mathbf{k}} q_{\mathbf{k}} \sin(\mathbf{k}\cdot\mathbf{g} + \tfrac{1}{4}\pi) \tag{3.1}$$

Equation (1.29) then takes the form:

$$i\hbar \frac{\partial b_{\mathbf{g}}}{\partial t} = \sum_{\mathbf{k}} \left[-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial q_{\mathbf{k}}^2} + \frac{M\omega_{\mathbf{k}}^2}{2} q_{\mathbf{k}}^2 - \left(\frac{2}{N}\right)^{1/2} A q_{\mathbf{k}} \sin(\mathbf{k}\cdot\mathbf{g} + \tfrac{1}{4}\pi) \right] b_{\mathbf{g}} - J \sum_{\mathbf{h}} e^{i\alpha \mathbf{g},\mathbf{g}+\mathbf{h}} b_{\mathbf{g}+\mathbf{h}} \tag{3.2}$$

where dispersion of the vibrational frequencies is given by:

$$\omega_k^2 = \omega_0^2 + \omega_1^2 \sum_{\mathbf{h}} \cos(\mathbf{k}\cdot\mathbf{h}), \tag{3.3}$$

the sum going over nearest neighbors.

Apart from the phase factors multiplying J and the obvious two-dimensional features of (3.2), the subsequent development is essentially the same as that given in II, equations (1)–(19). The zeroth order ($J = 0$) eigenstates and eigenvalues are:

$$b_{g_1,\cdots N_{\mathbf{k}}^{(1)}\cdots}(\mathbf{g}, \cdots q_{\mathbf{k}} \cdots) = \delta_{\mathbf{g},\mathbf{g}_1} \prod_{\mathbf{k}} \Phi_{N_k^{(1)}} \left[\left(\frac{M\omega_k}{\hbar}\right)^{1/2} (q_{\mathbf{k}} - q_{\mathbf{k}}^{(1)}) \right] \tag{3.4}$$

$$E_{\cdots N_{\mathbf{k}}^{(1)}\cdots} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} (N_{\mathbf{k}}^{(1)} + \tfrac{1}{2}) + E_b \tag{3.5}$$

respectively, where

$$q_{\mathbf{k}}^{(1)} = \left(\frac{2}{N}\right)^{1/2} \frac{A}{M\omega_{\mathbf{k}}^2} \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) \tag{3.6}$$

are the equilibrium normal mode coordinates,[20]

$$\Phi_N(z)$$

are normalized harmonic oscillator wave functions, and

$$E_{\mathrm{b}} = -\left(\frac{2}{N}\right) \sum_{\mathbf{k}} \frac{A^2}{2M\omega_{\mathbf{k}}^2} \sin^2(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) \tag{3.7}$$

is the polaron binding energy.

[20] These arise as a result of incorporating the linear electron-lattice interaction term in zeroth order. The electron is taken to be initially self-trapped on site $\mathbf{g}_1$.

The zeroth order eigenstates and eigenvalues correspond to the physical situation in which the electron is localized at a given site, $\mathbf{g}_1$, and the vibrational state of the system is specified by a set of quantum numbers, $N_{\mathbf{k}}^{(1)}$, giving the degree of excitation of each vibrational mode.[21] It is to be noted that these modes differ from the conventional modes in that their equilibrium displacements, about which the system oscillates, depends on the location of the electron (cf. (3.4) and (3.6)).

As in II, the total wave function of the system is expanded in the "local-site" representation (3.4). The matrix elements of the perturbation (i.e., the J-proportional terms of (3.2)) are found to be:

$$(\mathbf{g}, \cdots, N_{\mathbf{k}} \cdots \mid V \mid \mathbf{g}_1, \cdots N_{\mathbf{k}}^{(1)} \cdots) = \sum_{\mathbf{h}} \delta_{\mathbf{g},\mathbf{g}_1+\mathbf{h}} \times (-Je^{i\alpha_{\mathbf{g},\mathbf{g}_1}})$$

$$\times \prod_{\mathbf{k}} \left\{ \left[1 - \frac{4}{N} (N_{\mathbf{k}}^{(1)} + \tfrac{1}{2}) \gamma_{\mathbf{k}\cdot\mathbf{h}} \cos^2 \left(\mathbf{k}\cdot\left(\mathbf{g}_1 + \frac{\mathbf{h}}{2}\right) + \frac{\pi}{4} \right) \right] \delta_{N_k, N_k^{(1)}} \right. \tag{3.8}$$

$$\left. \mp \left[\left(\frac{8}{N}\right)^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}} \gamma_{\mathbf{k}\cdot\mathbf{h}}^{1/2} \cos \left(\mathbf{k}\cdot\left(\mathbf{g}_1 + \frac{\mathbf{h}}{2}\right) + \frac{\pi}{4} \right) \left(\frac{N_k^{(1)} + \tfrac{1}{2} \pm \tfrac{1}{2}}{2} \right)^{1/2} \right] \delta_{N_k, N_k^{(1)} \pm 1} \right\}$$

where

$$\gamma_{\mathbf{k}\cdot\mathbf{h}} = \frac{A^2}{2M\omega_{\mathbf{k}}{}^2 \hbar\omega_{\mathbf{k}}} [1 - \cos(\mathbf{k}\cdot\mathbf{h})] \tag{3.9}$$

and $\epsilon_{\mathbf{k}\cdot\mathbf{h}}$ is equal to ± 1 according to whether $\mathbf{k}\cdot\mathbf{h}(\gtrless)0$.

From their form, it is clear that the matrix elements (3.8) give rise to site-jump transitions ($\mathbf{g}^{(1)} \to \mathbf{g}^{(1)} + \mathbf{h}$). These will be used to calculate the (magnetic) field-dependent part of the transition rate by time-dependent perturbation theory. Before proceeding with the calculation, however, it is of importance to distinguish between the so-called *diagonal* transitions in which all the $N_{\mathbf{k}}$ remain unaltered, and the so-called *nondiagonal* transitions in which some of them change by ± 1. As shown in II, the nondiagonal transitions are dominant in the site-jump regime (as defined in the introduction and in II by the inequality $T > T_c$). However, within the framework of standard time-dependent perturbation theory, a certain formal difficulty arises. Namely, as a result of the exact conservation of the unperturbed energy, characteristic of the diagonal transitions, their contribution to the jump probability, in the site-jump regime, is not time independent, but contains terms which diverge linearly and quadratically[22] with time. Strictly speaking, a proper treatment of the diagonal

[21] The notation $\cdots N_k^{(1)} \cdots$ refers to the totality of vibrational quantum numbers corresponding to the electron being localized on site 1.

[22] The quadratic divergences arise from the fact that higher order processes are operative. Such terms do not occur in II.

transitions would require the introduction of collision-damping into the theory. However, as also noted in II, there exists a simple physical criterion for the diagonal transitions to be of subsidiary importance, without recourse to a formal damping theory. Namely, all that is required is that the probability for a diagonal transition to occur in a mean lifetime of a local state be small compared to unity. As discussed in II, this requirement is by definition satisfied in the site-jump regime.

From the above remarks, it is apparent that the formally divergent contributions of the diagonal transitions to the standard perturbation calculation of the jump rate have no physical meaning, and should therefore be discarded altogether. This subtraction recipe will hence be employed in the present treatment without further ado.

Proceeding with the calculation, let us first obtain a general expression for the transition rate due to the interference of a first and second order process (to terms linear in the magnetic field). At the initial time $t_i = -\infty$, only the initial state (1)[23] is occupied. At some later time t, the first order amplitude and second order amplitude (via intermediate state 2) for occupancy of the final state (3) are, respectively

$$\begin{aligned} a_3^{(1)} &= \frac{1}{i\hbar}\int_{-\infty}^{t} dt'\, V_{31}^{(H)}\, e^{i\omega_{31}t} \\ a_3^{(2)} &= -\frac{1}{\hbar^2}\int_{-\infty}^{t} dt'\, V_{32}^{(H)}\, e^{i\omega_{32}t'} \int_{-\infty}^{t'} dt''\, V_{21}^{(H)}\, e^{i\omega_{21}t''} \end{aligned} \tag{3.10}$$

where

$$\omega_{ij} = \sum_{\mathbf{k}} \omega_{\mathbf{k}}(N_{\mathbf{k}}^{(i)} - N_{\mathbf{k}}^{(j)}); \qquad (i, j) = 1, 2, 3$$

and

$$V_{ij}^{(H)} = e^{i\alpha_{ij}} V_{ij}$$

are the (magnetic) field-dependent matrix element given by (3.8).

The transition rate associated with the interference of these two amplitudes

[23] The *states* 1, 2 and 3 represent the following collection of quantum numbers

$$\begin{aligned} 1 &\to (\mathbf{g}N_1\,, \cdots {}^{(1)}_{k} \cdots) \\ 2 &\to (\mathbf{g}_2\,, \cdots N_k^{(2)} \cdots) \\ 3 &\to (\mathbf{g}_3\,, \cdots N_k^{(3)} \cdots) \end{aligned}$$

The sums over the intermediate and final states are taken over the $N_{\mathbf{k}}$'s only; the site variables are, of course, fixed.

is given by:

$$\frac{\partial}{\partial t}\left(|\, a_3^{(1)} + a_3^{(2)}\,|^2 - |\, a_3^{(1)}\,|^2 - |\, a_3^{(2)}\,|^2\right) = \frac{\partial}{\partial t}\sum_3 \left(a_3^{(1)*}\, a_3^{(2)}\right) + \text{c. c.} = \sum_3 \left[\frac{\partial a_3^{(1)*}}{\partial t}\, a_3^{(2)} + a_3^{(1)*}\,\frac{\partial a_3^{(2)}}{\partial t}\right] + \text{c. c.} \tag{3.11}$$

The calculation is simplified by the formal artifice of turning on the potential adiabatically, i.e., $V_{ij} \sim e^{st}$, $-\infty < t \leqq 0$. For this case, the integrals over t' and t'' are readily performed, giving (at $t = 0$):

$$a_3^{(1)} = -\frac{V_{31}^{(H)}}{\hbar}\,\frac{1}{\omega_{31} - is}, \qquad \frac{\partial_{a3}^{(1)}}{\partial t} = \frac{V_{31}^{(H)}}{i\hbar}$$

$$a_3^{(2)} = \frac{V_{32}^{(H)} V_{21}^{(H)}}{\hbar^2}\,\frac{1}{(\omega_{31} - i2s)(\omega_{21} - is)}, \qquad \frac{\partial a_3^{(2)}}{\partial t} = -\frac{1}{i\hbar^2}\,\frac{V_{32}^{(H)} V_{21}^{(H)}}{(\omega_{21} - is)}$$

Use is made of the above expressions (as well as their complex conjugates), and the fact that the field independent matrix elements are symmetric ($V_{ij} = V_{ji}$).

It should now be noted that, as far as the Hall effect is concerned, only that part of (3.11) which is linear in the magnetic field need be considered. Denoting this linear part by $w_3^{(H)}$, one finds, after some algebra, that:

$$w_3^{(H)} = \frac{\alpha_{13} + \alpha_{32} + \alpha_{21}}{\hbar^3} \sum_{2,3} V_{13}\, V_{32}\, V_{21}\, \frac{1}{\omega_{21} - is}\left[\frac{1}{\omega_{31} + is} - \frac{1}{\omega_{31} - i2s}\right] + \text{c.c.}$$

Using the standard identity

$$\lim_{s \to 0^+} \frac{1}{\omega - is} = P\left(\frac{1}{\omega}\right) + i\pi\delta(x)$$

one gets

$$w_3^{(H)} = \frac{\alpha_{13} + \alpha_{32} + \alpha_{21}}{\hbar^3} \sum_{2,3} V_{13}\, V_{32}\, V_{21} \cdot \left[-P\left(\frac{1}{\omega_{21}}\right) i2\pi\delta(\omega_{21}) + 2\pi^2\delta(\omega_{21})\delta(\omega_{31})\right] + \text{c.c.}$$

The first term in the square brackets represents the contribution of "virtual" transitions, in that the intermediate site energy is not conserved. Such contributions clearly cancel when adding the c.c. expression, and hence do not contribute[24] to order H. Finally, using the standard Fourier-integral represen-

[24] It should be emphasized that this conservation of intermediate site energy is essential for the explanation of the Hall effect in the three-site case. The reason is that, as previously

tation of the Dirac-delta function

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau e^{i\tau x},$$

$w_3^{(H)}$ takes the form:

$$w_3^{(H)} = -\frac{\alpha}{\hbar^3}\sum_{2,3}\int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2\, V_{13}\, V_{32}\, V_{21}\, e^{i\omega_{31}\tau_1} e^{i\omega_{21}\tau_2},$$

where $\alpha = \alpha_{32} + \alpha_{21} + \alpha_{13}$ is given by (2.7).

Substituting for the matrix elements from (3.8), we get:

$$\begin{aligned}
w_3^{(H)} = \alpha\frac{J^3}{\hbar^3}\sum_{\substack{\dots N_k^{(2)}\dots \\ \dots N_k^{(3)}\dots}}\int_{-\infty}^{\infty} d\tau_1\int_{-\infty}^{\infty} d\tau_2 \times \prod_k \Bigg[& e^{i\omega_k(N_k^{(3)}-N_k^{(1)})\tau_1} e^{i\omega_k(N_k^{(2)}-N_k^{(1)})\tau_2} \\
\times \Bigg\{ & \left[1 - \frac{4}{N}\left(N_k^{(3)} + \frac{1}{2}\right)\gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}\cos^2\left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4}\right)\right]\delta_{N_k^{(1)},N_k^{(3)}} \\
\mp & \left[\left(\frac{8}{N}\right)^{1/2}\epsilon_{\mathbf{k}\cdot\mathbf{h}_{31}}\gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}^{1/2}\cos\left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4}\right)\left(\frac{N_k^{(3)} + \tfrac{1}{2} \pm \tfrac{1}{2}}{2}\right)^{1/2}\right]\delta_{N_k^{(1)},N_k^{(3)}\pm 1}\Bigg\} \\
\times \Bigg\{ & \left[1 - \frac{4}{N}\left(N_k^{(2)} + \frac{1}{2}\right)\gamma_{\mathbf{k}\cdot\mathbf{h}_{32}}\cos^2\left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4}\right)\right]\delta_{N_k^{(2)},N_k^{(3)}} \\
\mp & \left[\left(\frac{8}{N}\right)^{1/2}\epsilon_{\mathbf{k}\cdot\mathbf{h}_{32}}\gamma_{\mathbf{k}\cdot\mathbf{h}_{32}}^{1/2}\cos\left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4}\right)\left(\frac{N_k^{(2)} + \tfrac{1}{2} \pm \tfrac{1}{2}}{2}\right)^{1/2}\right]\delta_{N_k^{(2)},N_k^{(3)}\pm 1}\Bigg\} \\
\times \Bigg\{ & \left[1 - \frac{4}{N}\left(N_k^{(2)} + \frac{1}{2}\right)\gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}\cos^2\left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2}\right) + \frac{\pi}{4}\right)\right]\delta_{N_k^{(1)},N_k^{(2)}} \\
\mp & \left[\left(\frac{8}{N}\right)^{1/2}\epsilon_{\mathbf{k}\cdot\mathbf{h}_{21}}\gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}^{1/2}\cos\left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2}\right) + \frac{\pi}{4}\right)\right. \\
& \left.\cdot\left(\frac{N_k^{(2)} + \tfrac{1}{2} \pm \tfrac{1}{2}}{2}\right)^{1/2}\right]\delta_{N_k^{(1)},N_k^{(2)}\pm 1}\Bigg\}\Bigg]
\end{aligned} \tag{3.12}$$

We have, then, a product of factors, each pertaining to a given mode, $\mathbf{k}$, with its corresponding frequency phase factors. The summations over the $N_\mathbf{k}$'s for

---

discussed, it is necessary for the second order amplitude (via the intermediate site) to be 90° out of phase with the direct amplitude in the absence of the magnetic field. For virtual occupation of site 2, however, the two amplitudes turn out to be in phase, so that the field-dependent contribution to the jump rate is $\sim H^2$ or higher and is of no importance with regard to the Hall effect.

each mode are readily performed, giving

$$
\begin{aligned}
w_3^{(H)} &= \alpha \frac{J^3}{\hbar^3} \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \\
&\times \prod_k \Bigg\{ \Bigg[1 - \frac{4}{N} \left(N_k^{(1)} + \frac{1}{2} \right) \Bigg(\gamma_{\mathbf{k}\cdot\mathbf{h}_{31}} \cos^2 \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \\
&+ \gamma_{\mathbf{k}\cdot\mathbf{h}_{32}} \cos^2 \left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4} \right) + \gamma_{\mathbf{k}\cdot\mathbf{h}_{31}} \cos^2 \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \Bigg) \Bigg] \\
&- \frac{8}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{32}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{32}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{21}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4} \right) \cos \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \\
&\cdot \left[\frac{N_k^{(1)}}{2} e^{-i\omega_k \tau_2} + \frac{N_k^{(1)} + 1}{2} e^{i\omega_k \tau_2} \right] \\
&+ \frac{8}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{31}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{21}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \cos \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \\
&\cdot \left[\frac{N_k^{(1)}}{2} e^{-i\omega_k(\tau_1+\tau_2)} + \frac{N_k^{(1)} + 1}{2} e^{i\omega_k(\tau_1+\tau_2)} \right] \\
&+ \frac{8}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{31}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{32}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{32}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4} \right) \\
&\cdot \left[\frac{N_k^{(1)}}{2} e^{-i\omega_k \tau_1} + \frac{N_k^{(1)} + 1}{2} e^{i\omega_k \tau_1} \right] \Bigg\}
\end{aligned}
\tag{3.13}
$$

The first term in each factor refers to the case for which the N_k's of a given mode are unaltered for all three transitions. The other three terms refer to situations in which one of the transitions is diagonal and the other two change by ± 1.[25]

Writing the exponentials as sine and cosine, we have:

$$
\frac{N_k^{(1)}}{2} e^{-ix} + \frac{N_k^{(1)} + 1}{2} e^{ix} = \left(N_k^{(1)} + \frac{1}{2} \right) \cos x + i \sin x
$$

Next, the thermal average of the initial distribution of $N_k^{(1)}$'s is taken. One has

$$
\langle N_k^{(1)} \rangle = (e^{\beta\hbar\omega_k} - 1)^{-1}, \qquad \beta = \frac{1}{\kappa T}
$$

[25] Cases in which two jumps are diagonal and one non-diagonal, or all three change by ± 1, are inconsistent with the Kronecker-delta restrictions. The minus sign in front of the second term is due to the fact the N_k must first increase by one unit, and then decrease by one, if the direct transition is diagonal ($N_k^{(3)} = N_k^{(1)}$). For the other two terms, N_k increases or decreases by one for both nondiagonal jumps.

or

$$\left\langle N_k^{(1)} + \frac{1}{2} \right\rangle = \frac{1}{2} \coth \frac{\beta\hbar\omega_k}{2}$$

which when substituted into (3.13), yields

$$\begin{aligned}
w_3^{(H)} = \alpha \frac{J^3}{\hbar^3} \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \exp \sum_k \Bigg\{ & -\frac{2}{N} \coth \frac{\beta\hbar\omega_k}{2} \\
& \cdot \Bigg[\gamma_{\mathbf{h}\cdot\mathbf{k}_{31}} \cos^2 \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \\
& + \gamma_{\mathbf{k}\cdot\mathbf{h}_{32}} \cos^2 \left(\frac{\mathbf{k}\cdot\mathbf{h}_{21}}{2} + \frac{\pi}{4} \right) + \gamma_{\mathbf{k}\cdot\mathbf{h}_{21}} \cos^2 \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \Bigg] \\
& - \frac{4}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{32}} \gamma_{\mathbf{h}\cdot\mathbf{k}_{32}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{21}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4} \right) \cos \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \\
& \times \left[\coth \frac{\beta\hbar\omega_k}{2} \cos \omega_k \tau_2 + i \sin \omega_k \tau_2 \right] \\
& + \frac{4}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{31}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{21}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{21}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \cos \left(\mathbf{k}\cdot\left(\mathbf{h}_{32} + \frac{\mathbf{h}_{21}}{2} \right) + \frac{\pi}{4} \right) \\
& \times \left[\coth \frac{\beta\hbar\omega_k}{2} \cos \omega_k(\tau_1 + \tau_2) + i \sin \omega_k(\tau_1 + \tau_2) \right] \\
& + \frac{4}{N} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{31}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{31}}^{1/2} \epsilon_{\mathbf{k}\cdot\mathbf{h}_{32}} \gamma_{\mathbf{k}\cdot\mathbf{h}_{32}}^{1/2} \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2} + \frac{\pi}{4} \right) \cos \left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2} + \frac{\pi}{4} \right) \\
& \times \left[\coth \frac{\beta\hbar\omega_k}{2} \cos \omega_k \tau_1 + i \sin \omega_k \tau_1 \right] \Bigg\}
\end{aligned} \tag{3.14}$$

At this stage, a considerable simplification can be effected with regard to the summation over $\mathbf{k}$ by taking advantage of the rotation-inversional properties of $\omega_{\mathbf{k}}$. According to (3.3), it is clear that:

$$\omega_{-\mathbf{k}} = \omega_{\mathbf{k}}$$

and that

$$\omega_{\mathbf{k}'} = \omega_{\mathbf{k}} ,$$

where $\mathbf{k}'$ is obtained from $\mathbf{k}$ by rotating by $\pm\pi/3$ (about an axis perpendicular to the triangular lattice).

The first equality enables one to combine terms of $\pm\mathbf{k}$ in the first term of the sum, and thereby replace[26] $\cos^2(\frac{1}{2}\mathbf{k}\cdot\mathbf{h} + \frac{1}{4}\pi)$ by $\frac{1}{2}$. Applying both the first and the second relation to the remaining three terms in the sum allows one

[26] See II, footnote 1.

to group together $\mathbf{k}$'s related by rotations of $\pm\frac{1}{3}\pi$. Let us illustrate this for the second term appearing in the exponent in (3.14). Writing this term out in full, we have

$$\sum_{\mathbf{k}} (-)\frac{4}{N}\left(\frac{A^2}{2M\omega_k{}^2\hbar\omega_k}\right)^2\left[\sin\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2}\sin\frac{\mathbf{k}\cdot\mathbf{h}_{21}}{2}\right]\left[\cos\left(\frac{\mathbf{k}\cdot\mathbf{h}_{32}}{2}+\frac{\pi}{4}\right)\right.$$
$$\left.\times\cos\left(\mathbf{k}\cdot\left(\mathbf{h}_{32}+\frac{\mathbf{h}_{21}}{2}\right)+\frac{\pi}{4}\right)\right]\left[\coth\frac{\beta\hbar\omega_k}{2}\cos\omega_k\tau_2+i\sin\omega_k\tau_2\right]$$

Applying the standard trigonometric identities for the product of two sine or cosine functions, and neglecting terms which are odd in $\mathbf{k}$ (and hence cancel in taking the sum), we obtain

$$\sum_{\mathbf{k}} (-)\frac{4}{N}\left(\frac{A^2}{2M\omega_{\mathbf{k}}{}^2\hbar\omega_{\mathbf{k}}}\right)\left[\cos\left(\frac{\mathbf{k}}{2}\cdot(\mathbf{h}_{32}-\mathbf{h}_{21})\right)-\cos\left(\frac{\mathbf{k}}{2}\cdot(\mathbf{h}_{32}+\mathbf{h}_{21})\right)\right]$$
$$\times\frac{1}{2}\cos\frac{\mathbf{k}\cdot\mathbf{h}_{31}}{2}\left[\coth\frac{\beta\hbar\omega_k}{2}\cos\omega_k\tau_2+i\sin\omega_k\tau_2\right]$$

On further applying these identities to each of the product of the two cosine functions, we have

$$\sum_{\mathbf{k}}\frac{4}{N}\left(\frac{A^2}{2M\omega_{\mathbf{k}}{}^2\hbar\omega_{\mathbf{k}}}\right)\frac{1}{4}\left[(1-\cos(\mathbf{k}\cdot\mathbf{h}_{32}))+(1-\cos(\mathbf{k}\cdot\mathbf{h}_{21}))\right.$$
$$\left.-(1-\cos(\mathbf{k}\cdot\mathbf{h}_{31}))\right]\left[\coth\frac{\beta\hbar\omega_k}{2}\cos\omega_k\tau_2+i\sin\omega_k\tau_2\right]$$

The individual sums involving $\mathbf{h}_{32}$, $\mathbf{h}_{21}$, and $\mathbf{h}_{31}$ (which constitute all possible relative site vectors connecting the three mutually near neighbor sites) are, in effect, equal by the above cited property of rotational symmetry. We finally get (cf. Eq. (3.9))

$$\sum_{\mathbf{k}}\frac{1}{N}\gamma_{\mathbf{k}\cdot\mathbf{h}}\left(\coth\frac{\beta\hbar\omega_k}{2}\cos\omega_k\tau_2+i\sin\omega_k\tau_2\right)$$

Treating the last two terms of (3.14) in a similar fashion, we obtain:

$$w_3^{(H)}=\alpha\frac{J^3}{\hbar^3}\exp\left(-\frac{3}{N}\sum_{\mathbf{k}}\gamma_{\mathbf{k}\cdot\mathbf{h}}\coth\frac{\beta\hbar\omega_k}{2}\right)$$
$$\cdot\int_{-\infty}^{\infty}d\tau_1\int_{-\infty}^{\infty}d\tau_2\exp\left[F(\tau_1)+F(\tau_2)+F(\tau_1+\tau_2)\right] \tag{3.15}$$

where

$$F(\tau)\equiv\sum_{\mathbf{k}}\frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\left(\coth\frac{\beta\hbar\omega_k}{2}\cos\omega_{\mathbf{k}}\tau+i\sin\omega_{\mathbf{k}}\tau\right) \tag{3.16}$$

Before evaluating the principal contributions, it is noted that (3.15) is not properly time-independent, because it includes diagonal contributions, as discussed at the beginning of this section. This may be demonstrated by the fact that $F(\tau)$, given by (3.16), can be made arbitrarily small for sufficiently large τ (for sufficient dispersion of the vibrational spectrum, as discussed at length in II). It can then be observed that (3.15) will diverge linearly or quadratically in time depending upon whether one or both variables of integration (i.e., τ_1 and τ_2) are taken to be large, respectively. The identification, physical interpretation and subtraction of the varieties of diagonal contributions are carried out in Appendix C. The final nondiagonal rate is given by (C.2), which is quoted here for convenience:

$$\begin{aligned} w_3^{(H)} = \alpha \frac{J^3}{\hbar^3} \exp\left\{\sum_{\mathbf{k}}\left[-\frac{3}{N}\gamma_{\mathbf{h}\cdot\mathbf{k}} \coth\frac{\beta\hbar\omega_k}{2}\right]\right\} \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \\ \times \{[\exp F(\tau_1) - 1][\exp F(\tau_2) - 1][\exp F(\tau_1 + \tau_2) - 1] \\ + [\exp F(\tau_1) - 1][\exp F(\tau_2) - 1] \\ + [\exp F(\tau_2) - 1][\exp F(\tau_1 + \tau_2) - 1] \\ + [\exp F(\tau_1 + \tau_2) - 1][\exp F(\tau_1) - 1]\} \end{aligned} \tag{C.2}$$

In contrast to (3.15), the integrand goes to zero (with sufficient rapidity) as $|\tau_1| \to \infty$, $|\tau_2| \to \infty$.

As discussed in II (bottom p. 359 et seq.), (C.2) can be evaluated in the two regimes where $|F(\tau)|$ is small or large compared to unity. For fixed values of the parameters A, M, ω_0, and ω_1 which determine $\gamma_{\mathbf{k}\cdot\mathbf{h}}$, these cases occur at low and high temperatures respectively. At low temperatures, the transitions are characterized by a series of processes in which one, two, three etc. phonons are either emitted or absorbed (single-phonon regime). Of interest for the present paper is the high-temperature quantum limit (multiphonon regime) for which[27]

$$\sum_{\mathbf{k}} \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \operatorname{csch}\frac{\beta\hbar\omega_{\mathbf{k}}}{2} \gg 1 \tag{3.17}$$

For this case, the method of steepest descents is appropriate and will now be employed.

Proceeding with the evaluation of (C.2) let us consider the first term of the integrand, consisting of a product of three factors. This integral is denoted by

[27] It will be later shown that it is (3.17), rather than the inequality

$$\textstyle\sum_{\mathbf{k}} (\gamma_{\mathbf{k}\cdot\mathbf{h}}/N) \coth (\beta\hbar\omega_k/2) \gg 1$$

suggested by (3.16), which is required for the applicability of the method of steepest descents.

I_1. The difficulty in performing the indicated integrations τ_1 and τ_2 are coupled in the terms $\cos \omega_k(\tau_1 + \tau_2)$, $\sin \omega_k(\tau_1 + \tau_2)$. We can decouple the variables and treat $\tau_3 = \tau_1 + \tau_2$ on an equal footing with τ_1 and τ_2 by multiplying the integrand by the Dirac delta function $\delta(\tau_1 + \tau_2 - \tau_3)$, which is here written in terms of its Fourier integral representation. This procedure gives

$$I_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\alpha \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \int_{-\infty}^{\infty} d\tau_3 [\exp F(\tau_1) - 1][\exp F(\tau_2) - 1] \cdot [\exp F(\tau_3) - 1] \exp [i\alpha(\tau_1 + \tau_2 - \tau_3)] \tag{3.18}$$

The integrations over τ_1, τ_2 and τ_3 are all independent. Consider first the integral over τ_1. Expressing $F(\tau_1)$ by means of (3.16), and introducing the transformation

$$\tau_1 = \frac{i\beta\hbar}{2} + \theta_1,$$

one obtains, after some trigonometric manipulation,

$$\int_{-\infty - i(\beta\hbar/2)}^{\infty - i(\beta\hbar/2)} d\theta_1 \left[\exp \left(\sum_k \frac{\gamma_{k \cdot h}}{N} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cos \omega_k \theta_1 \right) - 1 \right] \exp (i\alpha\theta_1).$$

As previously stated, the integral is to be evaluated by the method of steepest descents. As is well known, one must first determine the locations of the saddle points. Ignoring, for the present, the term -1 in the square bracket of the integrand, the saddle point, $\theta_1^{(0)}$, is given by:

$$\sum_k \frac{\gamma_{k \cdot h}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sin \omega_k \theta_1^{(0)} = i\alpha$$

In performing the integrations over τ_2 and τ_3, one introduces analogous transformations

$$\tau_2 = \frac{i\beta\hbar}{2} + \theta_2$$

$$\tau_3 = \frac{i\beta\hbar}{2} + \theta_3.$$

The integration over θ_2 proceeds just like that for θ_1; for θ_3, it is noted that α appears with a minus sign because of the form of the argument of the delta function. Accordingly, the saddle points $\theta_2^{(0)}$ and $\theta_3^{(0)}$ are given by:

$$\sum_k \frac{\gamma_{k \cdot h}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sin \omega_k \theta_2^{(0)} = i\alpha$$

$$\sum_k \frac{\gamma_{k \cdot h}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sin \omega_k \theta_3^{(0)} = -i\alpha.$$

By setting

$$\begin{aligned} \theta_1^{(0)} &= \theta_2^{(0)} = i\eta \\ \theta_3^{(0)} &= -i\eta, \end{aligned} \tag{3.19}$$

the saddle point equations can be written as the single relation

$$\sum_{\mathbf{k}} \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh \omega_k \eta = \alpha(\eta) \tag{3.20}$$

In principle, one would find η as a function of α and subsequently integrate over α. Since the transcendental equation (3.18) is not readily solvable, we instead regard α as a parametric function of the (as yet undetermined) saddle point, η. A subsequent integration over η will later locate the saddle points exactly. Eliminating α by means of (3.20), one obtains

$$\begin{aligned} I_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\eta \Bigg[\sum_{\mathbf{k}} \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k^2 \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cosh\omega_k\eta \Bigg] \\ \cdot \exp\left\{ -\frac{\beta\hbar}{2} \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh\omega_k\eta \right\} \\ \times \int_{-\infty-(i\beta\hbar/2)}^{\infty-(i\beta\hbar/2)} d\theta_1 \left[\exp\left\{ \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cos\omega_k\theta_1 \right\} - 1 \right] \\ \cdot \exp\left\{ i\theta_1 \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh\omega_k\eta \right\} \\ \times \int_{-\infty-(i\beta\hbar/2)}^{\infty-(i\beta\hbar/2)} d\theta_2 \left[\exp\left\{ \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{2} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cos\omega_k\theta_2 \right\} - 1 \right] \\ \cdot \exp\left\{ i\theta_2 \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh\omega_k\eta \right\} \\ \times \int_{-\infty-(i\beta\hbar/2)}^{\infty-(i\beta\hbar/2)} d\theta_3 \left[\exp\left\{ \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cos\omega_k\theta_3 \right\} - 1 \right] \\ \cdot \exp\left\{ -i\theta_3 \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh\omega_k\eta \right\} \end{aligned} \tag{3.21}$$

The one-dimensional integrals are evaluated in Appendix D by the method of steepest descents, where questions of contour, subsidiary saddle points, and behavior for large θ are considered. It is there argued that the principal saddle points are located at[28]

[28] This is identical with (3.19).

$$\mathrm{Re}[\theta_\alpha^{(0)}] = 0$$

$$\mathrm{Im}[\theta_\alpha^{(0)}] = \pm\eta$$

((+) is taken for $\alpha = 1, 2$; (−) for $\alpha = 3$).

Performing the integrations over the θ_α, we are left with the following integral over η:

$$I_1 = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\eta \left[\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\,\omega_k^{\,2}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\cosh\omega_k\eta\right] \cdot \left[\frac{\pi}{\tfrac{1}{2}\sum_k (\gamma_{\mathbf{k}\cdot\mathbf{h}}/N)\,\omega_k^{\,2}\,\mathrm{csch}\,(\beta\hbar\omega_k/2)\cosh\omega_k\eta}\right]^{3/2} \cdot \exp\left\{\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\left[3\cosh\omega_k\eta - 3\,(\omega_k\eta)\sinh\omega_k\eta - \frac{\beta\hbar\omega_k}{2}\sinh\omega_k\eta\right]\right\}.$$

The integrand is manifestly well behaved for large η. The saddle point for η is readily found to be:

$$\eta^{(0)} = -\frac{\beta\hbar}{6}.$$

Finally, carrying out the η integration by the method of steepest descents, we get,

$$I_1 = \exp\left(\frac{3}{N}\sum_k \gamma_{\mathbf{k}\cdot\mathbf{h}}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\cosh\frac{\beta\hbar\omega_k}{6}\right)\frac{2\pi}{\sqrt{3}}\left(\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\,\omega_k^{\,2}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\cosh\frac{\beta\hbar\omega_k}{6}\right)^{-1}$$

It remains to consider the last three terms of (C.2), consisting of products of two factors. Let us consider the second term, I_2. The individual integrals are formally similar to the integral required for the evaluation of the two-site jump rate (cf. II, (54) et seq.). They are readily carried out by making the change of variables

$$\tau_1 = \frac{i\beta\hbar}{2} + \tau_1'$$

$$\tau_2 = \frac{i\beta\hbar}{2} + \tau_2'$$

and carrying out the τ_1', τ_2' integrations by steepest descents. The result is

$$I_2 = \int_{-\infty}^{\infty} d\tau_1'\,[\exp F(\tau_1') - 1]\int_{-\infty}^{\infty} d\tau_2'\,[\exp F(\tau_2') - 1]$$

$$= 2\pi\left(\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\,\omega_k^{\,2}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\right)^{-1}\left(\exp\frac{2}{N}\sum_k \gamma_{\mathbf{k}\cdot\mathbf{h}}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\right).$$

By suitable changes of variable, the last two terms of (C.2) are clearly identical

with the above. Combining results, one has:

$$w_3^{(H)} = \alpha \frac{J^3}{\hbar^3}\left[\frac{2\pi}{\sqrt{3}} \exp\left(-\frac{6}{N}\sum_k \gamma_{k\cdot h} \operatorname{csch}\frac{\beta\hbar\omega_k}{2} \sinh\frac{\beta\hbar\omega_k}{3} \sinh\frac{\beta\hbar\omega_k}{6}\right)\right.$$
$$\times \left(\sum_k \frac{\gamma_{k\cdot h}}{N}\,\omega_k{}^2 \operatorname{csch}\frac{\beta\hbar\omega_k}{2} \cosh\frac{\beta\hbar\omega_k}{6}\right)^{-1}$$
$$+ 6\pi \exp\left(-\frac{1}{N}\sum_k \gamma_{k\cdot h} \operatorname{csch}\frac{\beta\hbar\omega_k}{2}\right) \exp\left(-\frac{3}{N}\sum_k \gamma_{k\cdot h} \tanh\frac{\beta\hbar\omega_k}{4}\right) \tag{3.22}$$
$$\left.\times \left(\sum \frac{\gamma_{k\cdot h}}{N}\,\omega_k{}^2 \operatorname{csch}\frac{\beta\hbar\omega_k}{2}\right)^{-1}\right]$$

where, to simplify the argument of the exponent of the first term, we have used the identity

$$\coth\frac{x}{2} - \operatorname{csch}\frac{x}{2}\cosh\frac{x}{6} \equiv 2\operatorname{csch}\frac{x}{2}\sinh\frac{x}{3}\sinh\frac{x}{6}.$$

It will now be demonstrated that the second term of (3.22) is very much less than the first in the high-temperature quantum limit (defined by (3.17)) of present interest.

With use of the identity

$$\tanh\frac{x}{4} \equiv \frac{\cosh(x/2) - 1}{\sinh(x/2)},$$

the ratio of the second term to the first can be written

$$\left[\frac{\sum_k(\gamma_{k\cdot h}/N)\,\omega_k{}^2 \operatorname{csch}(\beta\hbar\omega_k/2)\cosh(\beta\hbar\omega_k/6)}{\sum_k(\gamma_{k\cdot h}/N)\,\omega_k{}^2 \operatorname{csch}(\beta\hbar\omega_k/2)}\right]$$
$$\times \exp\left\{-\sum_k \frac{\gamma_{k\cdot h}}{N} \operatorname{csch}\frac{\beta\hbar\omega_k}{2}\left[3\left(\cosh\frac{\beta\hbar\omega_k}{6} - 1\right) + 1\right]\right\}$$

But since the quantity appearing in the square brackets is always greater than one, it follows from inequality (3.17) that this ratio is in fact very much less than unity. Hence, in what follows, the second term of (3.22) will be discarded.

The final result for the (magnetic) field-dependent jump rate in the high-temperature quantum limit is written in the form:

$$w_3^{(B)} = \left(\frac{e\mathbf{H}\cdot\mathbf{A}_{321}}{\hbar c}\right) w_0^{(0)}, \tag{3.23}$$

where

$$w_0^{(0)} = \frac{J^3}{\hbar^3}\frac{2\pi}{\sqrt{3}}\left(\sum_k \frac{\gamma_{k\cdot h}}{N}\omega_k^2 \operatorname{csch}\frac{\beta\hbar\omega_k}{2}\cosh\frac{\beta\hbar\omega_k}{6}\right)^{-1} \times \exp\left(-\frac{6}{N}\sum_k \gamma_{k\cdot h}\operatorname{csch}\frac{\beta\hbar\omega_k}{2}\sinh\frac{\beta\hbar\omega_k}{3}\sinh\frac{\beta\hbar\omega_k}{6}\right) \tag{3.24}$$

is the (electric and magnetic) field-independent part of $w_3^{(H)}$.

The physical significance of the various factors appearing in (3.24) is best seen in the classical high-temperature limit:

$$\beta\hbar\omega_k = \frac{\hbar\omega_k}{\kappa T} \ll 1$$

In this limit, the hyperbolic sine, cosecant, and cosine get replaced by their argument, reciprocal argument, and unity, respectively. Making use of (3.9), one finds, in this limit, that:

$$w_3^{(H)}(\mathrm{cl}) = \left(\frac{e\mathbf{H}\cdot\mathbf{A}_{321}}{\hbar c}\right) w_0^{(0)}(\mathrm{cl}), \tag{3.25}$$

where

$$w_0^{(0)}(\mathrm{cl}) = \frac{J^3}{\hbar}\frac{2\pi}{3\sqrt{3}\,\kappa T\,\epsilon_3} e^{-\epsilon_3/\kappa T}, \tag{3.26}$$

where

$$\epsilon_3 = \frac{1}{N}\sum_k \frac{A^2}{3M\omega_k^2}(1 - \cos\mathbf{k}\cdot\mathbf{h})$$

is the three-site activation energy (discussed in footnote 14 and Appendix B).

Thus, in this limit, (3.23) takes the form of a classical activation process, (3.25). This result is in fact identical with (2.19) obtained in the classical occurrence-probability approximation.

A discussion of the characteristic temperature variation of $w_3^{(H)}$ will be deferred to the next section where the calculation of the Hall effect is carried out explicitly for the three-site case.

IV. HALL COEFFICIENT CALCULATION FOR THE THREE SITE CONFIGURATION

Up to this point, consideration has been confined to the calculation of the magnetic field-dependent transition rate, $w^{(H)}(i \to k)$, associated with the elementary (two-dimensional) three-site configuration[29] shown in Fig. 1. In

[29] Such a configuration is intended as a prototype of structures possessing three mutually nearest neighbor sites (e.g., fcc). As we have seen, this geometrical property is of fundamental importance.

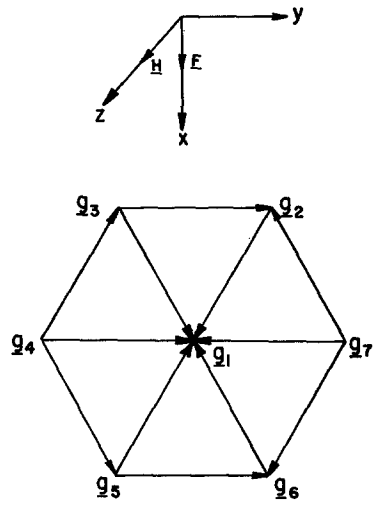

FIG. 2. Triangular lattice structure

order to calculate the Hall effect, one must compute the net drift velocity associated with the totality of field dependent jump rates connecting a given site to all its nearest neighbors. Such nearest neighbors form a regular hexagon as shown in Fig. 2. The electric field, **F**, is applied in the plane of the crystal in the x direction as shown; the magnetic field, **H**, is taken to be normal to the plane in the z direction.

Before proceeding with the calculation, it is important to point out a crucial feature responsible for the nonvanishing Hall effect: namely, the asymmetry associated with the magnetic field dependence. In particular, for a given direction of the field, the contribution to the field-dependent jump rate due to the interference process $(1 \rightarrow 2 + 1 \rightarrow 3 \rightarrow 2)$ is just the *negative* of that due to the (independent) process $(1 \rightarrow 3 + 1 \rightarrow 2 \rightarrow 3)$ considered in the last section. This is clear from symmetry considerations. This property will appear explicitly later in the calculation.

Proceeding, the net drift velocity of an electron initially located on site i ($i = 1$, see figure 2) is given by[30]

$$\mathbf{v} = \sum_k \mathbf{h}_{ik} w(i \rightarrow k)$$

where the $\mathbf{h}_{ik}$ are the totality of relative site vectors connecting the initial site $i = 1$ with all its nearest neighbors k, and where $w(i \rightarrow k)$ is the total jump rate to first order in H.

[30] The nearest neighbors sites (k) are assumed to be unoccupied.

It is convenient to write $\mathbf{v}$ as

$$\mathbf{v} = \mathbf{v}^{(0)} + \mathbf{v}^{(H)},$$

where

$$\mathbf{v}^{(0)} = \sum_k \mathbf{h}_{ik} w_F^{(0)}(i \to k) \tag{4.1}$$

$$\mathbf{v}^{(H)} \sum_{j,k} \mathbf{h}_{ik} w_F^{(H)}(i \to j \to k) \tag{4.2}$$

are the components of velocity independent of H and linear in H, respectively. Here, $w^{(0)}(i \to k)$ is given by[31]

$$w_F^{(0)}(i \to k) = \exp\left(\frac{e\mathbf{F}\cdot\mathbf{h}_{ki}}{2\kappa T}\right) w_2 \tag{4.3}$$

where w_2 is the elementary two-site jump rate given in II in the quantal and classical limits, respectively. These are quoted here for convenience:

$$w_2^{(\text{q.m.})} = \frac{J^2}{\hbar^2}\left[\frac{2\pi}{(1/N)\sum_k 2\gamma_k \omega_k^2 \operatorname{csch}(\beta\hbar\omega_k/2)}\right]^{1/2} \exp\left(-\frac{1}{N}\sum_k 2\gamma_k \tanh\frac{\beta\hbar\omega_k}{4}\right)$$

$$w_2^{(\text{cl.})} = \frac{J^2}{\hbar}\left[\frac{\pi}{4\kappa T\epsilon_2}\right]^{1/2} \exp\left(-\frac{\epsilon_2}{\kappa T}\right)$$

(ϵ_2 is the two-site activation energy discussed in Appendix B).

The (magnetic) field dependent jump rate $w^{(H)}(i \to j \to k)$ appearing in (4.2) is given by[32]

$$w_F^{(H)}(i \to j \to k) = \exp\left[\frac{e\mathbf{F}\cdot(\mathbf{h}_{ji} + \mathbf{h}_{ki})}{3\kappa T}\right] w_0^{(H)}(i \to j \to k) \tag{4.4}$$

where

$$w_0^{(H)}(i \to j \to k) = \left(\frac{eHA_{kji}}{hc}\right) w_0^{(0)}. \tag{4.5}$$

Here, $w_0^{(0)}$ is given by (3.24) and (3.26) in the quantum and classical limits, respectively. Further,

$$A_{kji} = \pm A, \tag{4.6}$$

where

$$A = \frac{\sqrt{3}}{4} a^2$$

[31] Both this and (4.4) are obvious generalizations of (2.20).

[32] The notation $(i \to j \to k)$ means that the second order process contributing to the transiton $(i \to k)$ occurs via intermediate site j.

is the (scalar) area of the elementary equilateral triangle defined by the mutually nearest neighbor sites, and

$(+)$ or $(-)$ is taken[33] according to whether the area $(i \to j \to k)$ is traversed in a counterclockwise or clockwise sense, respectively.

The calculation of $\mathbf{v}^{(0)}$, given by (4.1), will first be performed. To find the various $w^{(0)}(i \to j)$, one need only note the electric field dependence given by (4.3), obtaining

$$\begin{aligned} w^{(0)}(1 \to 2) &= w^{(0)}(1 \to 3) = e^{\chi} w_2 \\ w^{(0)}(1 \to 5) &= w^{(0)}(1 \to 6) = e^{-\chi} w_2 \\ w^{(0)}(1 \to 4) &= w^{(0)}(1 \to 7) = w_2 \end{aligned} \tag{4.7}$$

where

$$\chi = \frac{eF}{\kappa T} \frac{\sqrt{3}}{2} a.$$

From (4.1), one obtains (noting that the electric field is taken as infinitesimal, so that $\chi \ll 1$):

$$\begin{aligned} (v^{(0)})_x &= -\frac{3}{2} \frac{eF}{\kappa T} a^2 w_2 \\ (v^{(0)})_y &= 0 \end{aligned} \tag{4.8}$$

The drift mobility is then

$$\mu_D = \frac{v_x^{(0)}}{F} = \frac{3}{2} \frac{ea^2}{\kappa T} w_2 \tag{4.9}$$

In the high-temperature classical limit, one has

$$\mu_D^{(cl)} = \frac{3}{2} \frac{ea^2}{\kappa T} \frac{J^2}{\hbar} \left(\frac{\pi}{4\kappa T \epsilon_2} \right)^{1/2} e^{-\epsilon_2/\kappa T} \tag{4.10}$$

Next, $\mathbf{v}^{(H)}$, given by (4.2), will be calculated. In order to write down the various $w_F^{(H)}(i \to k \to j)$, one must take account of the change of sign of A_{kji} (given by (4.5)) which occurs with every interchange of two site indices. From (4.4), one obtains:

$$\begin{aligned} w_F^{(H)}(1 \to 2 \to 3) &= e^{\eta} w_0^{(H)} = -w_F^{(H)}(1 \to 3 \to 2) \\ w_F^{(H)}(1 \to 3 \to 4) &= e^{\eta/2} w_0^{(H)} = -w_F^{(H)}(1 \to 4 \to 3) \\ w_F^{(H)}(1 \to 4 \to 5) &= e^{-\eta/2} w_0^{(H)} = -w_F^{(H)}(1 \to 5 \to 4) \end{aligned}$$

[33] These signs apply for **H** directed *out* of the plane of the crystal, as shown in figure 2; for **H** directed into the plane, the signs would, of course, be reversed.

$$w_F^{(H)}(1 \to 5 \to 6) = e^{-\eta} w_0^{(H)} = -w_F^{(H)}(1 \to 6 \to 5)$$
$$w_F^{(H)}(1 \to 6 \to 7) = e^{-\eta/2} w_0^{(H)} = -w_F^{(H)}(1 \to 7 \to 6)$$
$$w_F^{(H)}(1 \to 7 \to 2) = e^{\eta/2} w_0^{(H)} = -w_F^{(H)}(1 \to 2 \to 7),$$

where $w_0^{(H)}$ represents the *absolute magnitude* of $w_0^{(H)}(1 \to j \to k)$ (given by (4.5)), and where

$$\eta = \frac{eF}{3\kappa T}\sqrt{3}a.$$

These relations explicitly show the fundamental asymmetry property cited at the beginning of this section.

Substituting these expressions into (4.2), again noting that $\eta \ll 1$, one finds, after some trigonometry and algebra, that[34]:

$$\begin{aligned} (v^{(H)})_x &= 0 \\ (v^{(H)})_y &= -\sqrt{3}\,\frac{eFa^2}{\kappa T}\, w_0^{(H)} \end{aligned} \tag{4.11}$$

The Hall angle, θ, is given by

$$\theta = \frac{j_y^{(H)}}{j_x^{(0)}},$$

where

$$\mathbf{j} = -ne\mathbf{v} = \mathbf{j}^{(0)} + \mathbf{j}^{(H)}$$

is the total electronic current. Hence, from (4.8) and (4.11), one obtains[35]

$$\theta = \frac{v_y^{(H)}}{v_x^{(0)}} = +\frac{2}{\sqrt{3}}\frac{w_0^{(H)}}{w_2}$$

The Hall mobility, μ_H, is

$$\mu_H = \frac{c}{H}\theta = +\frac{2}{\sqrt{3}}\frac{c}{H}\frac{w_0^{(H)}}{w_2} \tag{4.12}$$

In the classical limit, one finds that

$$\mu_H^{(cl)} = +\frac{2}{\sqrt{3}}\left(\frac{e}{\hbar}\frac{\sqrt{3}}{4}a^2\right) J\left[\frac{\pi}{4\kappa T\epsilon_2}\right]^{1/2} e^{-(\epsilon_3-\epsilon_2)/3\kappa T} \tag{4.13}$$

Finally, the (three-site) Hall coefficient, R_3, is given by

$$R_3 = -\frac{1}{nec}\frac{\mu_H}{\mu_D}$$

[34] This is properly in the direction of the classical Lorentz force, $[\mathbf{F} \times \mathbf{H}]$.

[35] The Hall angle (defined as the angle of deflection due to the magnetic field) is thus positive (i.e., counterclockwise, as in the classical case of a free-electron gas).

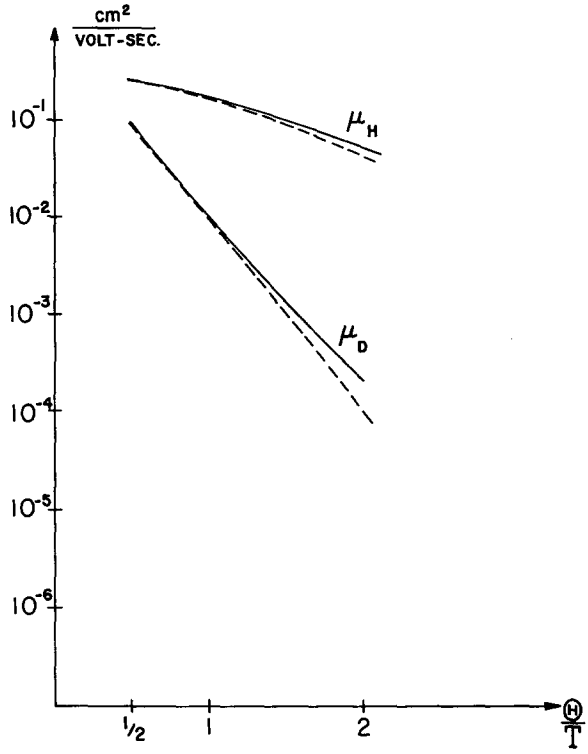

FIG. 3. Hall and drift mobilities vs. temperature for three-site configuration. (The dashed curves give the corresponding classical approximations.)

where n is the density of (excess) electrons. In the classical limit, one obtains

$$R_3^{(\mathrm{cl})} = -\frac{2}{3\sqrt{3}}\frac{1}{nec}\left(\frac{\kappa T}{J}\right)e^{(2\epsilon_2-\epsilon_3)/\kappa T} \tag{4.14}$$

In Fig. 3, we have plotted[36] the drift and Hall mobilities [(4.9) and (4.12)], respectively, as a function of (reciprocal) temperatures. The corresponding classical approximations [using (4.10) and (4.13)] are also shown by the dashed curves. It is first noted that the Hall mobility has a different activation energy than the drift mobility (specifically, it is ⅓ of the latter). This is simply a consequence of the fact that $\epsilon_3 = \frac{4}{3}\epsilon_2 < 2\epsilon_2$ as previously noted (see footnote 14). Further, the fact that $\mu_H > \mu_D$ is indicative of the fact that R_3 is *larger* than normal. Equation (4.13) shows that this is indeed the case; both of the factors

[36] This is based on representative values of the parameters given in II (footnote 17): $J = \hbar\omega_0 = 0.04$ ev, $\gamma = 10$, $\epsilon_2 = 5\hbar\omega_0 = 0.2$ ev, $a = 4$Å. In the figure, $\Theta = \hbar\omega_0/\kappa$ is the Debye theta.

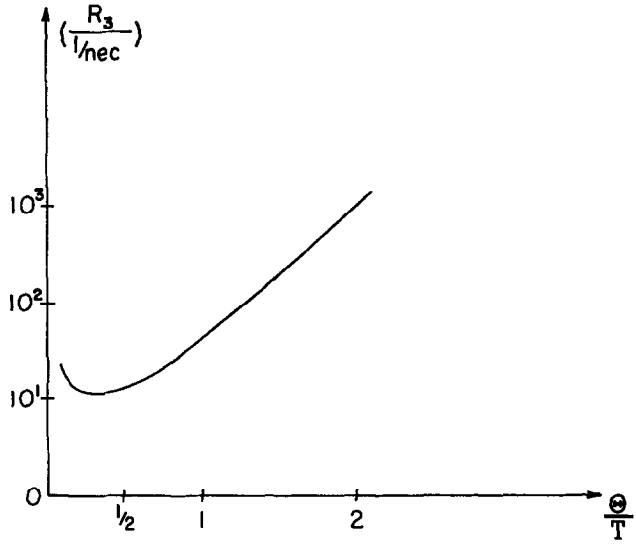

FIG. 4. Temperature dependence of the three-site Hall coefficient (classical approximation).

on the right hand side tend to make R_3 larger than normal. This feature is shown on Fig. 4, where we have[36] plotted the temperature dependence of R_3. Lastly, it may be noted that, as a result of the previously noted relationship between ϵ_2 and ϵ_3, R_3 is a *decreasing* function of temperature (except at very high temperatures, $T \gtrsim 3\ \Theta$).

V. FOUR-SITE HALL EFFECT (ORDER OF MAGNITUDE ESTIMATE)

The triangular lattice structure, which has just been treated, is of fundamental importance because of the relative simplicity of the associated interference mechanism, and the fact that its Hall effect is of lowest nonvanishing order in J (as we have seen, $R_3 \sim J^{-1}$). However, the calculation is not restricted to this geometry; of considerable importance is the Hall effect associated with a two-dimensional square (four-site) lattice array.[37] The calculation has not been carried out completely for this case[38]; hence only the essential features will be outlined. Such considerations are intended only to establish the order of magnitude of the four-site Hall coefficient relative to both the previously considered three-site Hall coefficient, and the "normal" value ($\sim -1/nec$).

[37] Such a geometry is representative of transition metal oxides of the NiO type to which this theory may find application.

[38] This will be done in a future publication. For a representative calculation, the reader is referred to III, Section IV-C-2.

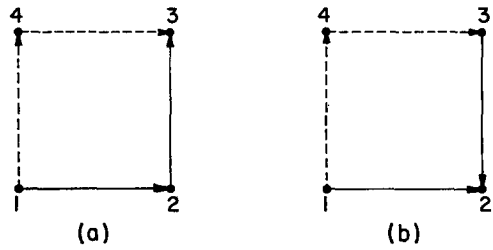

FIG. 5. Processes contributing to the four-site Hall effect. The dotted paths denote virtual transitions in which intermediate site energy is not conserved.

Figure 5 shows the two processes which appear to be of equal importance in contributing to the (magnetic) field dependent transition rate.

The first, (a), involves the interference of the two second order amplitudes[39] $(1 \rightarrow 2 \rightarrow 3 + 1 \rightarrow 4 \rightarrow 3)$. It turns out that in order for the two amplitudes to have the correct phase relationship (i.e., in order to get a linear contribution to the jump rate), it is necessary for three of the sites (1, 2, 3, or 1, 4, 3) to have a simultaneous coincidence in energy, while the energy of the other intermediate site (4 or 2, respectively) must differ from this common value. The second process, (b), consists of the interference of the third order amplitude $(1 \rightarrow 2 \rightarrow 3 \rightarrow 4)$, in which the energy of either site 2 or 3 is virtual, with the first order amplitude $(1 \rightarrow 4)$. In both cases, it is seen that the occurrence of a simultaneous coincidence in energy of *three* sites is a feature basic to the Hall effect (as in the three-site case).

On the basis of these considerations and the preliminary calculation given in III, the order of magnitude of the four-site Hall coefficient, R_4, relative to the three-site Hall coefficient, R_3, is given by:

$$R_4 \sim R_3 \left(\frac{J}{\Delta E}\right) \tag{5.1}$$

The energy denominator ΔE arises because of the virtual character of one of the intermediate sites as previously discussed; it turns out to be of the order of ϵ_2. The extra factor of J is simply due to the fact that four, rather than three, jumps are involved in the interference processes. For representative values of the parameters,[40] the magnitude of R_4 is comparable with the magnitude of the "normal" result, $-1/nec$.

VI. THE HALL EFFECT IN THE POLARON-BAND REGIME

As discussed in the introduction, the calculation of a Hall effect is of primary importance for the high temperature site-jump regime ($T > T_t$) which has just

[39] It is only this mechanism which is given preliminary consideration in III.

[40] See footnote 36.

been considered. This is because such a calculation theoretically establishes the existence of a non-vanishing Hall effect for this particular *hopping* mechanism of electronic conduction.

However, as was also pointed out in the introduction, for $T < T_t$, small polaron motion takes place in a band, the so-called "polaron-band," whose width is an exponentially decreasing function of temperature. The calculation of the Hall effect for this regime is considered to be of subsidiary importance, since it is formally similar to the case of an ordinary (tight-binding) electronic band. For this reason, and in the interest of not making the present paper too lengthy, this calculation[41] will be presented in detail in a later publication. For present purposes, consideration will be confined to some of the salient features of the band calculation, and a statement of the final results.

In the polaron-band regime, the appropriate zeroth order states are plane wave combinations of localized polaron states (3.4). The basic approach is to construct classical Bloch-type wave pockets from such plane wave states. As in the standard band theory, these wave pockets are found to propagate in k-space according to a classical Lorentz force equation. It is of particular interest in this connection that the effects of the basic "magnetic phase factors," $\alpha_{\mathbf{g},\mathbf{g}+\mathbf{h}}$, given by (1.24), turn out to be contained entirely in the conventional magnetic component $(e/c)[\mathbf{v}_\mathbf{k} \times \mathbf{H}]$ of the total force. Such a result is expected in the correspondence limit.

The Lorentz force term is then substituted into a Boltzmann equation describing the band motion. This equation is then solved[42] for the same three and four site structures considered in connection with the site-jump motion. Its solution is facilitated by a distinguishing feature of the present calculation, as compared to that for an ordinary electronic band: namely (as discussed in the introduction and in II), the extreme smallness (relative to the *electronic* bandwidth) of the *polaron* bandwidth

$$J_p = J \exp\left(-\frac{1}{N}\sum_k 2\gamma_{\mathbf{k}\cdot\mathbf{h}} \coth\frac{\beta\hbar\omega_k}{2}\right)$$

(This is the two-dimentional generalization of II (64)). In fact, J_p obeys the inequality

$$J_p \ll \kappa T$$

in contrast with the case normally encountered in semiconductors ($J \gtrsim \kappa T$). Consequently, it suffices to calculate the component of the current linear in H to lowest nonvanishing order in $J_p/\kappa T$. In so doing, it is found that

$$R_4 \sim -(1/nec)$$

[41] The calculation is given in III, Section V.

[42] This is done by an iterative procedure, treating the magnetic field as a perturbation.

but that

$$R_3 \sim -\left(\frac{\kappa T}{J_p}\right)\frac{1}{nec}$$

That is, the Hall effect associated with the three-site structure is larger than normal, just as in the site-jump regime.

Added Note

As was pointed out in footnote 35, the Hall angle is positive. A contrary result, obtained by one of the present authors (T. H.) in Section III of a paper on the Hall effect in impurity conduction (*8*), is wrong. The error arose from a mistake in vector algebra contained in equations (1.30) and (1.31) of that paper. As a result, the signs of all subsequent expressions containing the magnetic field should be altered. The conclusion is that, in the impurity case also, the sign of the Hall angle is positive.

APPENDIX A. CLASSICAL OCCURRENCE-PROBABILITY CALCULATION

In this appendix, the classical occurrence-probability for the three-site coincidence event, defined in the text proceeding (2.16), will be calculated.

According to this definition, this quantity is written[43]:

$$\begin{aligned} P_T{}^c(v_{31}, v_{12}, t_{12})\, dt_{31}\, dt_{12}\, dv_{31}\, dv_{12} \\ = Z^{-1} \int \cdots \int dx_1 \cdots dx_N\, dv_1 \cdots dv_N \exp\left\{-\frac{H_L + H_{int}}{\kappa T}\right\} \\ \times \delta[x_3(0) - x_1(0)]\, |\, v_{31}\, |\, dt_{31}\, \delta[v_3(0) - v_1(0) - v_{31}]\, dv_{31} \\ \times \delta[x_2(t_{12}) - x_1(t_{12})]\, |\, v_{12}\, |\, dt_{12}\, \delta[v_2(t_{12}) - v_1(t_{12}) - v_{12}]\, dv_{12}\,, \end{aligned} \tag{A.1}$$

where

$$Z = \int \cdots \int dx_1 \cdots dx_N\, dv_1 \cdots dv_N \exp\left(-\frac{H_L + H_{int}}{\kappa T}\right)$$

is the total partition function, and $H_L + H_{int} = H_L - Ax_1$ is the sum of the purely vibrational Hamiltonian and interaction Hamiltonian before[44] the transition.

One first transforms from the actual site displacements and velocities, x_i and v_i, to normal mode coordinates q_k and $\dot{q}_k$ via the orthogonal transformation:

$$x_i(t) = \left(\frac{2}{N}\right)^{1/2} \sum_{\mathbf{k}} \sin\,(\mathbf{k}\cdot\mathbf{g}_i + \tfrac{1}{4}\pi) q_k(t) \tag{A.2}$$

[43] A somewhat more careful definition of the analogous two-site occurrence-probability is given in (II, Appendix II) in terms of the unit step function.

[44] This feature has already been discussed in footnote 5.

In addition, it proves expedient to express the Kronecker-delta functions in terms of their Fourier-integral representation,

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\xi e^{-i\xi x}$$

Equation (A.1) then becomes:

$$\begin{aligned}
P_T^{\,c}(v_{12}, v_{31}, t_{12}) = {} & Z^{-1}\int\cdots\int\cdots dq_k\cdots d\dot{q}_k\cdots \\
& \cdot \exp\left\{-\sum_k\left[\frac{M\dot{q}_k^{\,2}}{2}+\frac{M\omega_k^{\,2}}{2}(q_k-q_k^{(1)})^2\right]\Big/\kappa T\right\} \\
& \times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\alpha\,\exp\left\{-i\alpha\left(\frac{2}{N}\right)^{1/2}\sum_k[\sin(\mathbf{k}\cdot\mathbf{g}_3+\tfrac{1}{4}\pi)\right. \\
& \qquad \left.-\sin(\mathbf{k}\cdot\mathbf{g}_1+\tfrac{1}{4}\pi)]q_k(0)\right\} \\
& \times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\beta\,\exp\left\{-i\beta\left[\left(\frac{2}{N}\right)^{1/2}\sum_k[\sin(\mathbf{k}\cdot\mathbf{g}_3+\tfrac{1}{4}\pi)\right.\right. \\
& \qquad \left.\left.-\sin(\mathbf{k}\cdot\mathbf{g}_1+\tfrac{1}{4}\pi)]\dot{q}_k(0)-v_{31}\right]\right\} \\
& \times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\gamma\,\exp\left\{-i\gamma\left(\frac{2}{N}\right)^{1/2}\sum_k[\sin(\mathbf{k}\cdot\mathbf{g}_1+\tfrac{1}{4}\pi)\right. \\
& \qquad \left.-\sin(\mathbf{k}\cdot\mathbf{g}_2+\tfrac{1}{4}\pi)]q_k(t_{12})\right\} \\
& \times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\delta\,\exp\left\{-i\delta\left[\left(\frac{2}{N}\right)^{1/2}\sum_k[\sin(\mathbf{k}\cdot\mathbf{g}_1+\tfrac{1}{4}\pi)\right.\right. \\
& \qquad \left.\left.-\sin(\mathbf{k}\cdot\mathbf{g}_2+\tfrac{1}{4}\pi)]\dot{q}_k(t_{12})-v_{12}\right]\right\}\times|v_{13}|\times|v_{12}|
\end{aligned} \tag{A.3}$$

where $q_k^{(1)} = (2/N)^{1/2}(A/M\omega_k^{\,2})\sin(\mathbf{k}\cdot\mathbf{g}_1+\tfrac{1}{4}\pi)$ are the equilibrium normal coordinates corresponding to the electron initially being trapped on site $\mathbf{g}_1$.

Introducing the explicit harmonic time dependence of the q_k, a further transformation is now performed from the q_k and $\dot{q}_k$ to the amplitudes and phases Q_k and δ_k, defined by:

$$\begin{aligned}
q_k(t) &= Q_k\cos(\omega_k t+\delta_k)+q_k^{(1)} \\
\dot{q}_k(t) &= -\omega_k Q_k\sin(\omega_k t+\delta_k)
\end{aligned} \tag{A.4}$$

The Jacobian for this transformation is $|J_k| = M\omega_k Q_k$. Equation (A.3)

then takes the form:

$$
\begin{aligned}
P_T{}^{c}(v_{12}, v_{31}, t_{12}) = Z^{-1} \int \cdots \int \cdots d\delta_k \cdots Q_k \cdot dQ_k \cdots \\
\times \exp\left\{-\sum_k \frac{M\omega_k{}^2 Q_k{}^2}{2\kappa T}\right\} \\
\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\alpha \exp\left\{-i\alpha\left(\frac{2}{N}\right)^{1/2}\sum_k [\sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi)\right. \\
\left. - \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi)][Q_k \cos\delta + q_k^{(1)}]\right\} \\
\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\beta \exp\left\{+i\beta\left(\frac{2}{N}\right)^{1/2}\sum_k [\sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi)\right. \\
\left. - \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi)]Q_k\,\omega_k \sin\delta + i\beta v_{31}\right\} \qquad \text{(A.5)} \\
\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\gamma \exp\left\{-i\gamma\left(\frac{2}{N}\right)^{1/2}\sum_k [\sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi)\right. \\
\left. - \sin(\mathbf{k}\cdot\mathbf{g}_2 + \tfrac{1}{4}\pi)][Q_k \cos(\omega_k t_{12} + \delta_k) + q_k^{(1)}]\right\} \\
\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\delta \exp\left\{i\delta\left(\frac{2}{N}\right)^{1/2}\sum_k [\sin(\mathbf{k}\cdot\mathbf{g}_2 + \tfrac{1}{4}\pi)\right. \\
\left. - \sin(\mathbf{k}\cdot\mathbf{g}_2 + \tfrac{1}{4}\pi)]Q_k\,\omega_k \sin(\omega_k t_{12} + \delta_k) + i\delta v_{12}\right\} \times |v_{12}| \times |v_{31}|.
\end{aligned}
$$

where

$$
Z = \int \cdots \int \cdots d\delta_k \cdots Q_k\, dQ_k \cdots \exp\left(-\sum_k \frac{M\omega_k{}^2 Q_k{}^2}{2\kappa T}\right).
$$

It is apparent that the integrations in (A.5) correspond to simple averages over the arbitrary phases, δ_k, and thermal averages over the amplitudes, Q_k.

First, the average over the phases is performed by grouping together the four exponential factors depending on δ_k. Taking advantage of the fact that each term in the sum over $\mathbf{k}$ is infinitesimal, $\sim(N^{-1/2})$, the exponentials are expanded to second order and the averages taken. Noting that the linear terms average to zero and that

$$
\langle \sin\delta_k \cos\delta_k \rangle_{\text{av}} = \langle \sin(\omega_k t_{12} + \delta_k) \cos(\omega_k t_{12} + \delta_k) \rangle_{\text{av}} = 0,
$$

the resulting expression[45] is readily integrated over the Q_k. One is left with the following[46]:

$$P_T{}^C(v_{12}, v_{31}, t_{12}) = \frac{1}{(2\pi)^4} \iiint_{-\infty}^{\infty} d\alpha\, d\beta\, d\gamma\, d\delta$$
$$\cdot \exp\{-[a_1(\alpha^2 + \gamma^2) + a_2(\beta^2 + \delta^2) + a_3\,\alpha\gamma - a_4(\alpha\delta + \beta\gamma) + a_5\,\beta\delta + ia_6(\alpha + \gamma) - iv_{31}\,\beta - iv_{12}\,\delta]\} \mid v_{12} \mid \cdot \mid v_{31} \mid \tag{A.6}$$

where

$$a_1 = \frac{1}{N}\sum_k \frac{\kappa T}{M\omega_k{}^2}[1 - \cos(\mathbf{k}\cdot\mathbf{h})] \qquad a_5 = \frac{1}{N}\sum_k \frac{\kappa T}{M}[1 - \cos(\mathbf{k}\cdot\mathbf{h}]\cos\omega_k t_{12}$$

$$a_2 = \frac{\kappa T}{M} \qquad a_6 = \frac{4\epsilon_2}{A}$$

$$a_3 = \frac{1}{N}\sum_k \frac{\kappa T}{M\omega_k{}^2}[1 - \cos(\mathbf{k}\cdot\mathbf{h})]\cos\omega_k t_{12}$$

$$a_4 = \frac{1}{N}\sum_k \frac{\kappa T}{M\omega_k}[1 - \cos(\mathbf{k}\cdot\mathbf{h})]\sin\omega_k t_{12}$$

At this point, some simplification of the integrations are possible for some special cases of interest.

Case 1. $t_{12} \to \infty$. This describes an event for which the coincidence of the initial and intermediate sites (1–2) is not correlated with that of the initial and final sites (1–3). As one would expect, this probability is just the product of the individual two-site occurrence-probabilities (given by II, (II–18)). Noting that

$$a_3 = a_4 = a_5 = 0$$

because of dispersion of the vibrational spectrum, as previously discussed, the integrations are readily performed, giving:

$$P_T{}^C(v_{12}, v_{31}) = \left(\frac{A}{2\pi}\right)^2 \left(\frac{M}{4\pi\kappa T}\right)\left(\frac{\pi}{4\kappa T\epsilon_2}\right)$$
$$\cdot \exp\left(-2\epsilon_2 - \frac{Mv_{31}^2}{4\kappa T} - \frac{Mv_{12}^2}{4\kappa T}\right) \mid v_{12} \mid \mid v_{31} \mid \tag{A.7}$$
$$= P_T{}^C(v_{31}) \cdot P_T{}^C(v_{12})$$

[45] The various relative site vectors gets replaced by a single $\mathbf{h}$, in effect, by taking advantage of the invariance of ω_k to rotation of $\mathbf{k}$ by $\pi/3$ as discussed in the text (see text material following (3.13)).

[46] In its present form, (A.6) is manifestly invariant to time reversal of the unperturbed lattice motions. That is, $v_{31} \to -v_{31}$, $v_{12} \to -v_{12}$, $t_{12} \to -t_{12}$ and a simultaneous change of variables $\beta \to -\beta$, $\delta \to -\delta$ give (A.6) back again.

as anticipated, where $P_T{}^C(v_r)$ is the two-site occurrence-probability given by (II, (II–18)), and ϵ_2 is the two-site activation energy discussed in the text (see footnote 14).

Case 2. $t_{12} = 0$. This describes a simultaneous triple coincidence event. Noting that

$$a_3 = a_1\,, \qquad a_4 = 0, \qquad a_5 = a_2\,,$$

one finds that

$$P_T{}^C(v_{31}\,, v_{12}\,, 0) = \frac{A^2}{2\pi}\frac{M}{2\pi\sqrt{3}\kappa T}\frac{1}{3\sqrt{3}\kappa T\epsilon_3} \cdot \exp\left[-\frac{\epsilon_3}{\kappa T} - \frac{M}{6\kappa T}(v_{12}^2 + v_{23}^2 + v_{31}^2)\right] | v_{12} | \, | v_{31} | \tag{A.8}$$

where $\epsilon_3 = \frac{4}{3}\epsilon_2$ is the three-site activation energy discussed in the text, and in Appendix B. Thus, the probability of a simultaneous coincidence of three sites is *larger* than the product of the individual two-site probabilities, representing a correlation of the latter when the two pair of coincidences occur at the same time.

Case 3. In the text ((2.17) et seq.) we had occasion to investigate $P_T{}^C$ for small t_{12}. This is most easily done for the final case to be examined, that of the Einstein spectrum, $\omega_k = \omega_0$. Expanding the trigonometric function for small $(\omega_0 t_{12}) \ll 1$, and retaining terms quadratic in this quantity, it is found that the probability function can be written as:

$$P_T{}^C(v_{31}\,, v_{12}\,, t_{12}) \sim \exp\left(-\frac{\epsilon_2}{\kappa T} - \frac{Mv_{31}^2}{4\kappa T} + \mu + \nu\right),$$

where

$$\mu = -\frac{\epsilon_2}{\kappa T}(1 + C_3\,\omega_0{}^2 t_{12}^2) - \frac{4}{3}\frac{Mv_{31}^2}{\kappa T}(\omega_0\, t_{12})^2 + \frac{Av_{31}\, t_{12}}{6\kappa T}(1 + C_4\,\omega_0{}^2 t_{12}^2),$$

$$\begin{aligned}\nu = {} & C_1\frac{\epsilon_2}{\kappa T}(\omega_0{}^2 t_{12}^2) \\ & - C_2\frac{(Av_{31}\, t_{12}/2)(1 + C_5\,\omega_0{}^2 t_{12}^2) - Av_{12}\, t_{12}(1 + C_6\,\omega_0{}^2 t_{12}^2)}{\kappa T} \\ & - \frac{Mv_{12}\, v_{31}}{3\kappa T}(1 + C_7\,\omega_0{}^2 t_{12}^2) - \frac{Mv_{31}^2/4}{3\kappa T}(1 + C_8\,\omega_0{}^2 t_{12}^2) - \frac{Mv_{12}^2}{3\kappa T}.\end{aligned} \tag{A.9}$$

where the C's are numerical coefficients of the order of unity.

It is noted that for $t_{12} = 0$,

$$\mu \rightarrow -\frac{\epsilon_2}{3\kappa T},$$

$$\nu \rightarrow -\frac{v_{12}\, v_{31}}{3\kappa T} - \frac{v_{31}^2/4}{3\kappa T} - \frac{v_{12}^2}{3\kappa T},$$

and (A.9) reduces correctly to (A.8).

APPENDIX B. IDENTIFICATION OF ACTIVATION ENERGIES WITH MINIMAL ENERGIES REQUIRED TO ESTABLISH COINCIDENCE CONFIGURATIONS

In this appendix, we shall supply a physical interpretation to the activation energies corresponding to the two and three site situations: namely, the *minimum* potential energy, in excess of the binding energy, required to establish the two and three site coincidence configurations, respectively. The importance of such configurations, as discussed in the test, is that the principal contributions to the net jump rate occur in their neighborhood.

Two-Site Case: Neglect of Dispersion

Consider a coincidence of sites $\mathbf{g}_1$ and $\mathbf{g}_3$. The electron is initially localized on site $\mathbf{g}_1$ whose equilibrium position is $x_1^{(0)} = \frac{A}{M\omega_0^2}$. The total potential energy is:

$$E = \frac{M\omega_0^2}{2}\left(x_1 - \frac{A}{M\omega_0^2}\right)^2 + \frac{M\omega_0^2}{2}\, x_3^2 + \sum_{m\neq 1,3} \frac{M\omega_0^2}{2}\, x_m^2 + E_b \qquad \text{(B.1)}$$

subject to the restriction:

$$x_1 = x_3 \qquad \text{(B.2)}$$

Alternately, we introduce the Lagrange multiplier, λ, and minimize the quantity $[(E - E_b) + \lambda(x_1 - x_3)]$ subject to no restrictions. This leads to the relation[47]:

$$0 = M\omega_0^2\left(x_1^{(c)} - \frac{A}{M\omega_0^2} - \lambda\right)\delta x_1 + M\omega_0^2(x_3^{(c)} + \lambda)\,\delta x_3 \qquad \text{(B.3)}$$

Choosing $\lambda = x_3^{(c)}$, δx_1 is an independent variation, and $x_1^{(c)} = (A/M\omega_0^2) + \lambda$. From (B.2), we find

$$x_1^{(c)} = x_3^{(c)} = \frac{A}{2M\omega_0^2}; \qquad E^{(\min)} - E_b = \frac{A}{4M\omega_0^2} = \epsilon_2 \qquad \text{(B.4)}$$

[47] We distinguish the equilibrium displacement corresponding to the bound state, $x^{(0)}$, from the displacement corresponding to a coincidence event, $x^{(c)}$.

Initially, then, $x_1^{(0)} = A/M\omega_0^2$; it decreases to $x_1^{(c)} = \frac{1}{2}x_1^{(0)} = A/2M\omega_0^2$, while x_3 increases from zero to meet it, establishing the coincidence.

INCLUSION OF DISPERSION (TWO SITES)

Our restriction (B.2) takes the form

$$\left(\frac{2}{N}\right)^{1/2} \sum_k q_k \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) = \left(\frac{2}{N}\right)^{1/2} \sum_k q_k \sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi) \quad \text{(B.5)}$$

The total energy with dispersion (in excess of the binding energy) is:

$$E - E_b = \sum_k \frac{M\omega_k^2}{2}\left[q_k - \left(\frac{2}{N}\right)^{1/2} \frac{A}{M\omega_k^2} \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi)\right]^2 \quad \text{(B.6)}$$

As before, the condition $\delta[(E - E_b) + \lambda(x_1 - x_3)] = 0$ leads to the relation:

$$0 = \sum_k M\omega_k^2 \left\{q_k^{(c)} - \left(\frac{2}{N}\right)^{1/2} \frac{A}{M\omega_k^2} \cdot[(1-\lambda)\sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) + \lambda \sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi)]\right\} \quad \text{(B.7)}$$

By the usual argument of Lagrange multipliers, we find:

$$q_k^{(c)} = \left(\frac{2}{N}\right)^{1/2} \frac{A}{M\omega_k^2}[(1-\lambda)\sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) + \lambda \sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi)] \quad \text{(B.8)}$$

Evaluating λ by means of condition (B.5), we get:

$$\lambda = \tfrac{1}{2} \quad \text{(B.9)}$$

Inserting into (B.8), and (B.6), we get:

$$E^{(\min)} - E_b = \sum_k \frac{M\omega_k^2}{2}\left[q_k^{(c)} - \frac{A}{M\omega_k^2}\left(\frac{2}{N}\right)^{1/2} \sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi)\right]^2 \quad \text{(B.10)}$$

We find, using $\mathbf{g}_1 = \mathbf{g}_3 + \mathbf{h}_{31}$, that:

$$E^{(\min)} - E_b = \sum_k \frac{A^2}{4NM\omega_k^2}(1 - \cos \mathbf{k}\cdot\mathbf{h}) = \epsilon_2 \quad \text{(B.11)}$$

THREE SITES: NO DISPERSION

In this case we write

$$E - E_b = \frac{M\omega_0^2}{2}\left(x_1 - \frac{A}{M\omega_0^2}\right)^2 + \frac{M\omega_0^2}{2}x_2^2 + \frac{M\omega_0^2}{2}x_3^2 + \sum_{m\neq 1,2,3} x_m^2 \quad \text{(B.12)}$$

subject to

$$x_1 = x_2 = x_3\,, \quad \text{(B.13)}$$

two independent conditions. Introducing two Lagrange multipliers, α and β, we obtain:

$$0 = \left[\left(x_1 - \frac{A}{M\omega_0^2}\right) + \alpha + \beta\right]\delta x_1 + [x_2 - \alpha]\delta x_2 + [x_3 - \beta]\delta x_3 \quad \text{(B.14)}$$

This gives

$$x_1^{(c)} = \frac{A}{M\omega_0^2} - \alpha - \beta; \qquad x_2^{(c)} = \alpha; \qquad x_3^{(c)} = \beta \quad \text{(B.15)}$$

Imposing (B.13), we find that

$$x_1^{(c)} = x_2^{(c)} = x_3^{(c)} = \frac{A}{3M\omega_0^2} \quad \text{(B.16)}$$

and

$$E^{(\min)} - E_b = \frac{A^2}{3M\omega_0^2} = \epsilon_3 = \frac{4}{3}\epsilon_2 \quad \text{(B.17)}$$

In this case, site $\mathbf{g}_1$ decreases to $\frac{2}{3}$ its initial value, while $\mathbf{g}_2$ and $\mathbf{g}_3$ increase from their zero equilibrium values to meet it, establishing the triple coincidence.

Including Dispersion (Three Sites)

In this case, our minimum condition reads:

$$\delta[(E - E_b) + \lambda_1(x_1 - x_3) + \lambda_2(x_1 - x_2)] = 0 \quad \text{(B.18)}$$

By the same procedures as before, the corresponding normal mode coordinates are found to be:

$$q_k^{(c)} = \left(\frac{2}{N}\right)^{1/2} \frac{A}{M\omega_0^2}[(1 - \lambda_1 - \lambda_2)\sin(\mathbf{k}\cdot\mathbf{g}_1 + \tfrac{1}{4}\pi) + \lambda_1 \sin(\mathbf{k}\cdot\mathbf{g}_2 + \tfrac{1}{4}\pi) + \lambda_2 \sin(\mathbf{k}\cdot\mathbf{g}_3 + \tfrac{1}{4}\pi)] \quad \text{(B.19)}$$

Using (B.13) expressed in terms of the normal mode coordinates, we have two equations which enable us to solve for λ_1 and λ_2. We find:

$$\lambda_1 = \lambda_2 = \tfrac{1}{3} \quad \text{(B.20)}$$

Using the relations among the relative site vectors, and rotational symmetry, we find, in the presence of dispersion.[48]

[48] Though generally true for arbitrary lattice geometries in the case of zero vibrational coupling, this simple numerical relationship applies to the case of non-vanishing coupling only for the three mutually near-neighbor situation of principal interest in this paper. For other geometries (e.g., square lattice array), the above argument of rotational symmetry cannot be applied, and the above result would not be expected to hold.

$$E^{(\min)} - E_b = \sum_{\mathbf{k}} \frac{A^2}{3NM\omega_{\mathbf{k}}^2}(1 - \cos \mathbf{k}\cdot\mathbf{h}) = \epsilon_3 = \frac{4}{3}\epsilon_2 \qquad \text{(B.21)}$$

APPENDIX C. SUBTRACTION OF DIAGONAL TRANSITIONS

In the text, it has been noted that (3.15) cannot represent a transition rate in the conventional sense, since it is not independent of t for large times. It is the purpose of this appendix to identify the varieties of diagonal transitions which are responsible for these divergences, and to subtract them from the total transition rate, as discussed in the text ((3.8) et seq.).

It is first noted that the part of (3.15) representing the totally diagonal transitions is the exponential factor independent of τ_1 and τ_2, namely,

$$\exp\left[-(3/N)\sum_k \gamma_{\mathbf{k}\cdot\mathbf{h}} \coth(\beta\hbar\omega_k/2)\right].$$

This is because the time dependent exponentials, $\exp[i(N_k^{(3)} - N_k^{(1)})\tau_1]$ and $\exp[i(N_k^{(2)} - N_k^{(1)})\tau_2]$ are all unity for such transitions. The removal of the quadratic divergence, then, is effected by subtracting just this term, i.e.,

$$\exp\left(-\frac{3}{N}\sum_k \coth\frac{\beta\hbar\omega_k}{2}\gamma_{\mathbf{k}\cdot\mathbf{h}}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau_1\, d\tau_2 \exp[F(\tau_1) + F(\tau_2) + F(\tau_1 + \tau_2)]$$

$$\longrightarrow \exp\left(-\frac{3}{N}\sum_k \coth\frac{\beta\hbar\omega_k}{2}\gamma_{\mathbf{k}\cdot\mathbf{h}}\right)$$

$$\cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\tau_1\, d\tau_2[\exp[F(\tau_1) + F(\tau_2) + F(\tau_1 + \tau_2)] - 1]$$

However, there remain additional divergences. For example, for $\tau_2 \to \infty$ only, the sum over $\mathbf{k}$ of the terms involving τ_2 and $(\tau_1 + \tau_2)$ become small, the integrand becomes essentially independent of τ_2, and we are led to a linear divergence. Recalling that the variable τ_2 applies to the initial-intermediate transitions $(1 \to 2)$, such a situation corresponds to transitions for which $N_k^{(3)} = N_k^{(1)} \pm 1$, for one or more modes, while $(\cdots N_k^{(2)} \cdots) = (\cdots N_k^{(1)} \cdots)$. The linear divergence is then due to the fact that the initial-intermediate transition is entirely diagonal.

Linear divergences are obtained for the two remaining cases in which the jumps $1 \to 2$ and $2 \to 3$ are diagonal. These latter three cases include, of necessity, the totally diagonal part as well as "partially diagonal" contributions. Hence, in performing the subtractions, $+3$ must be added in order that the totally diagonal part be subtracted once and not four times.

Performing all the subtractions, we obtain the following expression which is easily verified to be well behaved at large τ_1 and/or τ_2 :

$$w_3^{(H)} = \alpha \frac{J^3}{\hbar^3} \exp\left[\sum_{\mathbf{k}} - \left(\frac{3}{N}\right) \gamma_{k\cdot h} \coth \frac{\beta\hbar\omega_k}{2}\right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\tau_1 \, d\tau_2 \times \{\exp [F(\tau_1) + F(\tau_2) + F(\tau_1 + \tau_2)] - \exp F(\tau_1) - \exp F(\tau_2) - \exp F(\tau_1 + \tau_2) + 2\} \tag{C.1}$$

where $F(\tau)$ is given by (3.16).

By factorization, (C.1) is recast into the following more convenient form, as can be readily verified:

$$w_3^{(H)} = \alpha \frac{J^3}{\hbar^3} \exp\left\{-\frac{3}{N} \sum_k \gamma_{k\cdot h} \coth \frac{\beta\hbar\omega_k}{2}\right\} \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \times \{[\exp F(\tau_1) - 1][\exp F(\tau_2) - 1][\exp F(\tau_1 + \tau_2) - 1] + [\exp F(\tau_1) - 1][\exp F(\tau_2) - 1] + [\exp F(\tau_2) - 1][\exp F(\tau_1 + \tau_2) - 1] + [\exp F(\tau_1 + \tau_2) - 1][\exp F(\tau_1) - 1]\} \tag{C.2}$$

APPENDIX D. EVALUATION OF ONE-DIMENSIONAL INTEGRALS

In this appendix, the one-dimensional integrals appearing in (3.21) will be evaluated by the method of steepest descents. These integrals are required for the calculation of the quantum mechanical transition rate, (3.23).

A typical integral[49] is of the form

$$I = \int_{-\infty-(i\beta\hbar/2)}^{\infty-(i\beta\hbar/2)} d\theta \left[\exp\left(\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cos \omega_k \theta\right) - 1\right] \cdot \exp\left(i\theta \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sinh \omega_k \eta\right) \tag{D.1}$$

It is to be established that:

(1) Only the contribution at the "principal" saddle point $\theta^{(0)} = i\eta$ (to be derived later) need be considered, the contributions at the other "subsidiary" saddle points being small on account of dispersion of the vibrational spectrum.

(2) The integration (in the complex θ-plane) can be extended from $-\infty - i\beta\hbar/2$ to $\infty - i\beta\hbar/2$ by appropriately deforming the steepest descents contour.

The (-1) appearing in the integrand represents the subtraction of diagonal transitions as previously discussed. Its effect is to give an integrable contribution for large θ, and, hence, will be neglected until the behavior at large θ is considered.

[49] This applies to the integrals over θ_1 and θ_2 ; the integral over $\theta_3 (\eta \to -\eta)$ can be handled analogously.

The saddle points are found, of course, by setting the derivative of the argument of the exponent (D.1) equal to zero. This gives simply,

$$\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} (-\sin \omega_k \theta^{(0)} + i \sinh \omega_k \eta) = 0 \tag{D.2}$$

Writing θ in terms of its real and imaginary parts, i.e.,

$$\theta = \theta_r + i\theta_i ,$$

the real and imaginary parts of (D.2) become:

$$\text{Real part:} \quad \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \sin \omega_k \theta_r^{(0)} \cosh \omega_k \theta_i^{(0)} = 0 \tag{D.3}$$

$$\text{Imaginary part:} \quad \sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2} \cdot [-\cos \omega_k \theta_r^{(0)} \sinh \omega_k \theta_i^{(0)} + \sinh \omega_k \eta] = 0 \tag{D.4}$$

The "principal" saddle point occurs at

$$\theta_r^{(0)} = 0 \tag{D.5}$$

which clearly satisfies (D.3). From (D.4), it follows that

$$\theta_i^{(0)} = \eta \tag{D.6}$$

However, there may be other choices $\theta_r^{(0)} \neq 0$ for which the sums appearing in (D.3) and (D.4) vanish. These represent the "subsidiary" sadle points. At these points, we wish to examine the contributions to the real part of the argument of the (D.1), given by:

$$\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \operatorname{csch} \frac{\beta\hbar\omega_k}{2} [\cos \omega_k \theta_r^{(0)} \cosh \omega_k \theta_i^{(0)} - (\sinh \omega_k \eta)\omega_k \theta_i^{(0)}] \tag{D.7}$$

To show that such contributions are small, it is only necessary to point out that $\theta_i^{(0)}$ is a mildly increasing function of $\theta_r^{(0)}$; in particular, it will be later shown to be a logarithmic function of $\theta_r^{(0)}$ in the asymptotic (large θ) region. It then follows, from inequality (3.17) and for sufficient dispersion of the vibrational spectrum ($\omega_1 \sim \omega_0$) that (D.7) is much smaller at these points than at the principal saddle point (given by (D.5) and (D.6)).

It is instructive to consider the steepest descents contour sketched in Fig. 6. The dashed curve shows the contour applying to the narrow-band approximation[50] ($\omega_k \approx \omega_0$); the full curve includes vibrational dispersion.

[50] This case is not of physical interest, as discussed at length in II; it is only illustrative. From (D.3) and (D.4), the saddle points are found to be $\theta_i^{(0)} = 2\pi n/\omega_0$, $\theta_i^{(0)} = \eta$ (all n). The contour is found from the condition that the imaginary part of the argument of the

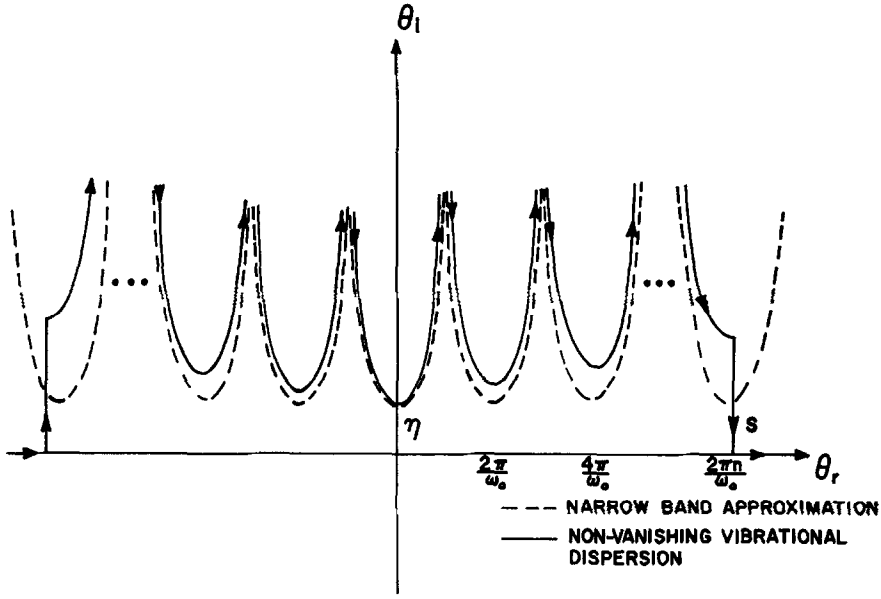

FIG. 6. Steepest descents contour

It is now to be demonstrated that the steepest descents contour can be taken to a subsidiary saddle point of sufficiently large θ_r, that the contribution from such a saddle point to $-i\beta\hbar/2$ (the path denoted by S in Fig. 6) may be made negligibly small.

What is called for is an examination of the argument of the integral (D.1) in the asymptotic (large θ) limit. This obtained by the method of steepest descents, and is analogous to the result II (61) obtained in the one-dimensional case. A preliminary examination gives the following:

$$\left(\exp A_{\mathbf{k}_0} \frac{\cos \omega_{\mathbf{k}_0} \Theta}{\omega_1^2 \theta} - 1\right) \exp\left(i\theta \sum_k A_k \sinh \omega_k \eta\right) \tag{D.8}$$

where

$$A_k = \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N} \omega_k \operatorname{csch} \frac{\beta\hbar\omega_k}{2}$$

$$A_{k_0} = \gamma_{\mathbf{k}_0\cdot\mathbf{h}} \omega_{k_0} \operatorname{csch} \frac{\beta\hbar\omega_{k_0}}{2}$$

---

exponential integrand be constant and pass through the appropriate saddle point, giving

$$\sinh \omega_k \theta_i = \sinh \omega_k \eta \left[\frac{\omega_0(\theta - 2\pi/\omega_0)}{\sin \omega_0(\theta - 2\pi n/\omega_0)}\right].$$

and

$$\omega_{k_0} = (\omega_0^2 - 2\omega_1^2)^{1/2}$$

is the mimimal lattice frequency.[51] Since $\theta_i^{(0)} \ll \theta_r^{(0)}$, as will be verified shortly, one can write

$$\frac{1}{\theta} = \frac{1}{\theta_r} - i\frac{\theta_1}{\theta_1^2} \cong \frac{1}{\theta_r}.$$

Then the first term of (D.8) can be written as

$$\exp[\phi(\theta_r, \theta_i) + i\psi(\theta_r, \theta_i)]$$

where

$$\phi(\theta_r, \theta_i) = A_{k_0}\frac{\cos\omega_{k_0}\theta_r \cosh\omega_{k_0}\theta_i}{\omega_1^2\theta_r} - \theta_i\sum_k A_k \sinh\omega_k\eta$$

$$\psi(\theta_r, \theta_i) = -A_{k_0}\frac{\sin\omega_{k_0}\theta_r \sinh\omega_k\theta_i}{\omega_1^2\theta_r} + \theta_r\sum_k A_k \sinh\omega_k\eta$$

To find the saddle points, we set $\partial\phi/\partial\theta_r = \partial\phi/\partial\theta_i = 0$, obtaining

$$\theta_r^{(0)} = 2\pi n/\omega_{k_0}$$

$$\sinh\omega_{k_0}\theta_i^{(0)} = \frac{\sum_k A_k \sinh\omega_k\eta}{A_{k_0}}\left(\frac{\omega_1^2}{\omega_{k_0}^2}\right)2\pi n$$

For sufficiently large n, the right hand side of the above equation can be made large compared to unity. In such a case, noting that $\sinh\omega_k\theta_i^{(0)} \cong \cosh\omega_k\theta_i^{(0)} \cong \frac{1}{2}\exp(\omega_{k_0}\theta_i^{(0)})$, one has

$$\theta_i^{(0)} = \frac{1}{\omega_{k_0}}\operatorname{Log}\left\{\frac{\sum_k A_k \sinh\omega_k\eta}{A_{k_0}}\frac{\omega_1^2}{\omega_{k_0}^2}(\omega_{k_0}\theta_r^{(0)})\right\}.$$

which verifies that $\theta_i^{(0)}$ is a logarithmic function of $\theta_r^{(0)}$ in the asymptotic region. The contribution at such a saddle point, moreover, is

$$\exp\phi(\theta_r^{(0)}, \theta_i^{(0)}) = \exp\left\{\frac{1}{\omega_{k_0}}\sum_k A_k \sinh\omega_k\eta\right\} \cdot\left[2\frac{\sum_k A_k \sinh\omega_k\eta}{A_{k_0}}\frac{\omega_1^2}{\omega_{k_0}^2}2\pi n\right]^{-\frac{1}{\omega_{k_0}}\sum A_k\sinh\omega_k\eta}$$

which is arbitrarily small for sufficiently large n. Further, where S intersects

[51] These occur at the midpoints of the boundaries of the hexagonal zone structure, appropriate to the triangular site configuration.

$-i\beta\hbar/2$, one has

$$\exp \phi\left(\theta_r^{(0)}, -\frac{i\beta\hbar}{2}\right) = \exp\left(\frac{A_{k_0}}{2\pi n}\frac{\omega_{k_0}}{\omega_1^2}\cosh\frac{\beta\hbar\omega_k}{2} + \frac{i\beta\hbar}{2}\sum_k A_k \sinh\omega_k\eta\right),$$

arbitrarily close to 1 in absolute magnitude.

Hence, along S and along the remaining contour $2\pi n/\omega_{k_0} \leqq \theta_r < \infty$, $\theta_i = -\beta\hbar/2$, the exponent of (D.1) can be expanded to first order, giving a small, integrable contribution. Identical considerations apply to the negative real axis.

The final result of these arguments is that only the contribution at the principal saddle point $\theta_r^{(0)} = 0$, $\theta_i^{(0)} = \eta$ need be taken. The result is:

$$I = \left[\frac{2\pi}{\sum_k (\gamma_{\mathbf{k}\cdot\mathbf{h}}/N)\,\mathrm{csch}\,(\beta\hbar\omega_k/2)\,\omega_k^2\cosh\omega_k\eta}\right]^{1/2} \times \exp\left[\sum_k \frac{\gamma_{\mathbf{k}\cdot\mathbf{h}}}{N}\,\mathrm{csch}\,\frac{\beta\hbar\omega_k}{2}\cosh\omega_k\eta - (\omega_k\eta)\sinh\omega_k\eta\right]$$

APPENDIX E. GAUGE INVARIANCE

It was pointed out in Section I that the fundamental property of the "magnetic phase factors" (1.25), though based on the use of a particular (symmetrical) gauge, can be obtained by an alternate method in which the gauge is arbitrary. It is the purpose of this appendix to demonstrate this explicitly.

An arbitrary vector potential is written as the sum of the vector potential in the symmetrical gauge plus the gradient of an arbitrary scalar function,

$$\mathbf{A}(\mathbf{r}) = \tfrac{1}{2}[\mathbf{H}\times\mathbf{r}] + \mathrm{grad}_{\mathbf{r}}\,\eta(\mathbf{r})$$

It proves convenient to rewrite this as the sum of a *local* vector potential (in the symmetrical gauge) plus the gradient of scalar function which is site dependent, i.e.

$$A(\mathbf{r}) = \tfrac{1}{2}[\mathbf{H}\times(\mathbf{r}-\mathbf{g})] + \mathrm{grad}_{\mathbf{r}}\,\eta_{\mathbf{g}}(\mathbf{r}) \tag{E.1}$$

where

$$\mathrm{grad}_{\mathbf{r}}\,\eta_{\mathbf{g}}(\mathbf{r}) = \mathrm{grad}_{\mathbf{r}}\,\eta(\mathbf{r}) + \tfrac{1}{2}[\mathbf{H}\times\mathbf{g}] \tag{E.2}$$

The local electronic wave functions, $\phi_{\mathbf{g}}(\mathbf{r})$, are the solution of (1.15) of Section I, except that, in the present case, $\mathbf{A}(\mathbf{r})$ is arbitrary and is given by (E.1).

Instead of the gauge transformation (1.16), however, one employs the alternate transformation

$$\Phi_{\mathbf{g}}(\mathbf{r}) = \exp\left\{-i\frac{e}{\hbar c}\eta_g(\mathbf{r})\right\}\chi(\mathbf{r}-\mathbf{g}) \tag{E.3}$$

where $\eta_{\mathbf{g}}(\mathbf{r})$ is defined by (E.2).

The $\chi(\mathbf{r}-\mathbf{g})$ then satisfy the same equations as in section 1, namely

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\,\mathrm{grad}_{\mathbf{r}}+\frac{e}{c}\frac{1}{2}\mathbf{H}\times(\mathbf{r}-\mathbf{g})\right)^2+U(\mathbf{r}-\mathbf{g})\right]\chi(\mathbf{r}-\mathbf{g})=E\,\chi(\mathbf{r}-\mathbf{g})$$

As argued in Section I (1.18) and following text), to first order in H (which is all that is of interest for the theory of the present paper), the $\chi(\mathbf{r}-\mathbf{g})$ are identical with the zero-field function $\phi(\mathbf{r}-\mathbf{g})$. Then, (E.3) becomes

$$\phi_g(\mathbf{r})=\exp\left\{-i\frac{e}{\hbar c}\eta_g(\mathbf{r})\right\}\phi(\mathbf{r}-\mathbf{g}) \tag{E.4}$$

The dependence of the $J_{\mathbf{g},\mathbf{g}+\mathbf{h}}$ on the magnetic field may be determined as before by substituting (E.4) into (1.14). This gives

$$J_{\mathbf{g},\mathbf{g}+\mathbf{h}}=-\int\phi(\mathbf{r}-\mathbf{g})\phi(\mathbf{r}-\mathbf{g}-\mathbf{h})U(\mathbf{r}-\mathbf{g})\exp\left\{-i\frac{e}{\hbar c}(\eta_{\mathbf{g}}(\mathbf{r})-\eta_{\mathbf{g}+\mathbf{h}}(\mathbf{r}))\right\}dV$$

Integrating (E.2) (applied to sites $\mathbf{g}$ and $(\mathbf{g}+\mathbf{h})$), and subtracting, one obtains

$$\eta_{\mathbf{g}}(\mathbf{r})-\eta_{\mathbf{g}+\mathbf{h}}(\mathbf{r})=\tfrac{1}{2}[\mathbf{H}\times\mathbf{h}]\cdot\mathbf{r}+[\eta_{\mathbf{g}}(0)-\eta_{\mathbf{g}+\mathbf{h}}(0)]$$

where $\eta_{\mathbf{g}}(0)$ and $\eta_{\mathbf{g}+\mathbf{h}}(0)$ are arbitrary scalar functions of $\mathbf{g}$ and $\mathbf{g}+\mathbf{h}$, respectively.

One then has

$$J_{\mathbf{g},\mathbf{g}+\mathbf{h}}=-\exp\left\{-i\frac{e}{\hbar c}[\eta_{\mathbf{g}}(0)-\eta_{\mathbf{g}+\mathbf{h}}(0)\right\}$$
$$\cdot\int\phi(\mathbf{r}-\mathbf{g})\phi(\mathbf{r}-\mathbf{g}-\mathbf{h})U(\mathbf{r}-\mathbf{g})\exp\{-ie[\mathbf{H}\times\mathbf{h}]\cdot\mathbf{r}/2\hbar c\}\,dV$$

Except for the exponential factor multiplying the integral, this result is just that obtained in Section 1.

To complete the argument, one only has to note that the final expressions for the jump probabilities are proportional to the sum of the "magnetic" phase factors around a closed path [e.g., (2.6), (3.13)]. However, the additional phase differences obtained in the present treatment, which are of the form

$$[\eta_{\mathbf{g}}(0)-\eta_{g'}(0)],$$

sum to zero around an arbitrary closed path. To illustrate this for our three-site geometry:

$$[\eta_2(0)-\eta_1(0)]+[\eta_3(0)-\eta_2(0)]+[\eta_1(0)-\eta_3(0)]\equiv 0$$

This established the fundamental property (1.25) for an arbitrary gauge.

RECEIVED: July 6, 1962

REFERENCES

1. T. HOLSTEIN, *Ann. Phys. (N.Y.)* **8,** 325, 343 (1959), to be referred to hereafter as I and II, respectively.
2. S. I. PEKAR, "Untersuchungen Über die Electronentheorie der Kristalle." Akademie-Verlag, Berlin, 1954.
3. G. R. ALLOCK, *Advanc. Phys.* **5,** 412 (1956).
4. F. J. MORIN, *Phys. Rev.* **93,** 1195 (1954).
5. R. R. HEIKES AND W. D. JOHNSTON, *J. Chem. Phys.* **26,** 582 (1957).
6. L. FRIEDMAN, Thesis (1961), University of Pittsburgh (unpublished), to be hereafter referred to as III.
7. G. ZILBERMAN, *Soviet Phys—JETP,* **2,** 650 (1956).
8. T. HOLSTEIN, *Phys. Rev.* **124,** 1329 (1961).