Occupancy-correlation corrections in hopping

Paul Gartner and Radu Pitiş
Theoretical Physics Group, Institute of Physics and Technology of Materials, P. O. Box MG-7, Bucharest-Măgurele R76900, Romania
(Received 23 October 1991)

A theory of correlated hopping diffusion on lattices with site-exclusion interaction is presented. We develop a variational procedure to estimate the diffusion constant D. We derive a method of obtaining several sequences of decreasing upper bounds for D. The method is extensively discussed for the particular sequence that starts with the mean-field approximant D^{MF}. The first term of a second sequence $\tilde{D}^{MF}$ is obtained by a suitable renormalization of the jump rates and is closer to D than D^{MF}. This modified mean-field approximation enables us to formulate an improved version of the Miller-Abrahams equivalence to a resistor network that includes correlation effects.

I. INTRODUCTION

Hopping transport on a lattice is treated in many cases by considering that the moving particles are noninteracting,[1-4] the problem being thus reduced, in a one-particle description, to that of a random walker. This simple picture is altered when one takes into account different types of interactions. One of the most important is the site-blocking interaction, which rules out the double occupancy of a given site.[5-12] This arises from the Pauli exclusion principle in the case of hopping fermions (electrons) (see Ref. 13 and references cited therein) and from a short-range (hard-core) repulsion in ionic conductors (see Ref. 14 and references cited therein). Even this most simple, shortest-range interaction brings about dynamic correlations which determine substantial changes in the transport properties of the lattice gas.

The one-particle description is usually restored by the mean-field (MF) approximation in which the dynamic environment encountered by the carrier is replaced by its equilibrium average. In this approximation the problem becomes exactly soluble in the case of periodic lattices[3] and one-dimensional chains,[4] where the MF value of the diffusion coefficient D^{MF} is readily obtained. It was in the frame of the MF approximation that Miller and Abrahams[1] proved the well-known equivalence between the dc conductivity of a hopping system and that of a resistor network. This approach was frequently used for obtaining approximate results and for proving different properties of the conductivity.

The evaluation of the exact diffusion constant D (or, using the Einstein relation, of the dc conductivity) is a many-body problem and thus a much more complicated one. An attempt to go beyond MF was made by Fedders and Sankey[5] who developed a systematic method based on diagrams for computing terms in the moment expansion of the correlation functions. The general formalism of Ref. 5 was further used by Fedders[7,8] for the study of two particular lattices.

A remarkable result was that of Richards[6] who obtained the exact value of $\sigma(0)$ for a one-dimensional (1D) periodic model with two kinds of sites, using a stationary-flow formalism. The same author made a generalization to arbitrary two-sublattice structures[9] and then obtained the frequency-dependent conductivity $\sigma(\omega)$ for the 1D model.[10] The latter was again discussed by Fedders and Richards.[11]

An interesting approach to the problem of occupancy correlation is due to Chase and Thouless.[13] Taking advantage of the formal similarity between the MF rate equation and the exact evolution, given by the master equation, they transposed the Miller-Abrahams argument from the real space onto the configuration space. Thus they obtained a "many-particle" resistor network with 2^N nodes, for a lattice of N sites. Even for 1D models the exponential increase of the computational volume with the sample length restricted the calculations to small systems and approximate (percolative) considerations.

In a previous paper[15] we treated the correlated hopping problem on a 1D random system with two types of sites. Corrections to the MF conductivity due to correlations were calculated using a stationary-flow approach.

Our present work reexamines the problem for the purpose of calculating the diffusion constant of identical particles (also called the constant of collective or of chemical diffusion, to distinguish it from the tracer diffusion constant). We make use of a transitory regime approach, in contrast to the steady-state one used in earlier papers which calculate the dc conductivity.[6,9,15] The two methods are equivalent.[16] Their results may be compared due to the Einstein relation between D and $\sigma(0)$.

In Sec. II the master equation theory of the diffusion constant is presented. It is shown that D is expressed by the resolvent of the master evolution operator A, restricted to the "one-particle" subspace S_1. The MF is identified as the approximation of restricting the operator *before* taking the resolvent. General properties of A are also discussed.

In Sec. III we examine the particular case of periodic lattices, which are of interest in their own right but also can be regarded as quite general large systems with periodic boundary conditions. Therefore we are interested in properties which do not depend on the size of the unit cell. It is shown that D can be identified from the

small frequency and quasimomentum asymptotic behavior of a certain diagonal element of the resolvent of A. Using the recursion method[17] we obtain a continued-fraction expansion for this matrix element, wherefrom by truncations at different levels we get a decreasing sequence of upper bounds for D. It is shown that one can improve the procedure by minimizing in a certain class the first term of the sequence. It is proved that in this case the first term is equal to D^{MF}.

Several applications of this scheme are presented in Sec. IV. They are not limited to periodic lattices but include general 1D chains too, illustrated on the particular case of the random binary chain.

In Sec. V we present an alternative, variational formulation of the general theory of Secs. II and III. This viewpoint provides additional freedom for the choice of the first step in the approximation sequence. The main result of this section can be stated as follows: we define an improved MF-type approximation which involves renormalized transition rates on the same lattice. The formal structure of the MF theory is left intact but due to the inclusion of correlation effects in the transition rates the modified diffusion constant $\tilde{D}^{\mathrm{MF}}$ is lower than the first correction to D^{MF} in the previous scheme, still remaining an upper bound for the exact value D. For lattices without triangles (i.e., in which the particle cannot return to its initial site at the third jump) the renormalization of the transition rates takes a particularly simple form. Adapting the Miller-Abrahams procedure to this approximation the renormalized resistors are easy to calculate. The obvious advantage over the network of Chase and Thouless is the maintaining of the original N-node lattice. As in Sec. III, $\tilde{D}^{\mathrm{MF}}$ is only the first term of a descending sequence of upper bounds for D. In the improved MF frame it is shown that correlations have a drastic effect on the low-temperature behavior of the conductivity, as illustrated on 1D binary chains.

Section VI contains a short summary.

II. MASTER EQUATION THEORY OF THE DIFFUSION CONSTANT

We denote by $\mathbf{x}_i$ and ε_i the positions and energies of the localized states on which hopping takes place and by $\mathcal{W}_{ji}$ the transition rate from $\mathbf{x}_i$ to $\mathbf{x}_j$. These quantities completely define the hopping system. The transition rates are non-negative quantities obeying the detailed balance relation ($\beta=1/kT$)

$$\mathcal{W}_{ij}e^{-\beta\varepsilon_j}=\mathcal{W}_{ji}e^{-\beta\varepsilon_i} . \tag{2.1}$$

A state of the system is characterized by a set of occupation numbers $n=\{n_i\}_i$, where n_i is either 0 or 1. The energy associated with the state n is given by

$$E_n=\sum_i \varepsilon_i n_i . \tag{2.2}$$

The master equation describes the evolution of the probability $P_n(t)$ of finding the system in the state n at time t. It reads[2]

$$\frac{d}{dt}P_n(t)=\sum_{n'}[\mathcal{W}_{nn'}P_{n'}(t)-\mathcal{W}_{n'n}P_n(t)] =-\sum_{n'}\mathcal{A}_{nn'}P_{n'}(t) . \tag{2.3}$$

$\mathcal{W}_{nn'}$ is the rate of the transition $n'\to n$. If we consider only individual transitions (i.e., one electron hops at a time) we have

$$\mathcal{W}_{nn'}=\sum_{k,l}\mathcal{W}_{kl}\delta_{n,n'+\delta_k-\delta_l}(1-n'_k)n'_l , \tag{2.4}$$

where δ_k is the set of occupation numbers $\delta_k=\{\delta_{k,i}\}_i$, the Kronecker δ in Eq. (2.4) means that n is generated from n' by transferring one electron from $\mathbf{x}_l$ to $\mathbf{x}_k$, and $\mathcal{W}_{kl}$ is the rate of this individual transition. The factor $(1-n'_k)n'_l$ forbids the transition unless $\mathbf{x}_l$ is occupied and $\mathbf{x}_k$ is empty in the initial state n'. Equations (2.4) and (2.2) ensure that $\mathcal{W}_{nn'}$ obey in their turn the detailed balance condition

$$\mathcal{W}_{nn'}e^{-\beta E_{n'}}=\mathcal{W}_{n'n}e^{-\beta E_n} \tag{2.5}$$

and therefore the macrocanonical distribution

$$P_n^0=\frac{1}{Z}e^{-\beta E_n} \tag{2.6}$$

(with the chemical potential μ taken as the energy zero) is the equilibrium solution of Eq. (2.3). Equations (2.2) and (2.4) show that the only interaction between the hopping particles considered here is the Pauli (occupancy, hard-core) interaction.

The (time-dependent) average of any n-dependent function F is

$$\bar{F}(t)=\sum_n F_n P_n(t) . \tag{2.7}$$

In the diffusion problem we are interested in the time evolution of the departure of the average occupation number $\bar{n}_i(t)$ from its equilibrium value $f_i=f(\varepsilon_i)$, f being the usual Fermi function. Denoting

$$\delta n_i=n_i-f_i , \tag{2.8}$$

one immediately gets from Eq. (2.3) the following equation:

$$\frac{d}{dt}\overline{\delta n_i}=\sum_j W_{ij}\left[\frac{\overline{\delta n_j}}{f_j(1-f_j)}-\frac{\overline{\delta n_i}}{f_i(1-f_i)}\right] -\sum_j W_{ij}\frac{f_i-f_j}{f_i(1-f_i)f_j(1-f_j)}\overline{\delta n_i\delta n_j} , \tag{2.9}$$

where W_{ij} are the symmetrized transition rates [see Eq. (2.1)]

$$W_{ij}=\mathcal{W}_{ij}(1-f_i)f_j=W_{ji} . \tag{2.10}$$

In order to get a closed equation for $\overline{\delta n_i}$ one usually neglects the second sum in Eq. (2.9). In this way one obtains the rate-equation description, or the mean-field (MF) approximation, on which most of the hopping-theory results are based.

If the lattice is uniform, i.e., all energies are equal, the

neglected term cancels out and the MF result is exact. In nonuniform lattices the approximation is good for small concentrations of electrons or holes but otherwise serious deviations are expected, especially at low temperatures,[6,13,15] and corrections to the MF approximation are needed. This is the occupancy correlation problem.

We begin with some definitions and notations. The master evolution operator $\mathcal{A}$, defined in Eq. (2.3), is given explicitly as

$$\mathcal{A}_{nn'}=\delta_{nn'}\sum_{n''}\mathcal{W}_{n''n}-\mathcal{W}_{nn'} . \tag{2.11}$$

In the space of the n-dependent functions, where it acts, we denote by $(\ ,\ )$ the canonical scalar product

$$(F,G)=\sum_n F_n^* G_n \tag{2.12}$$

and by $\mathcal{A}^\dagger$ the Hermitian conjugate of $\mathcal{A}$ with respect to it. The formal solution of Eq. (2.3) is

$$P(t)=e^{-\mathcal{A}t}P(0) \tag{2.13}$$

and for any (real-valued) F_n we have

$$\begin{aligned}\bar{F}(t)&=\sum_{nn'}F_n(e^{-\mathcal{A}t})_{nn'}P_{n'}(0)\\&=(F,e^{-\mathcal{A}t}P(0))=(e^{-\mathcal{A}^\dagger t}F,P(0)) .\end{aligned} \tag{2.14}$$

Therefore the time dependence can be carried over the "observables" by defining

$$F(t)=e^{-\mathcal{A}^\dagger t}F . \tag{2.15}$$

We define a new scalar product $\langle\ ,\ \rangle$ with the help of P^0

$$\langle F,G\rangle=(F,P^0G)=\sum_n F_n^* P_n^0 G_n , \tag{2.16}$$

in which, due to the detailed balance relation Eq. (2.5), the operator $\mathcal{A}^\dagger$ is Hermitic. Therefore, in order to avoid confusing notations we denote from now on $\mathcal{A}^\dagger$ by A. Moreover, it can be shown that $A\geq 0$, which accounts for the dissipative nature of the master evolution.

The diffusion constant tensor D in our hopping system is given by the $s\to 0$ limit $(s>0)$ of

$$\mathbf{e}D(s)\mathbf{e}=\frac{s^2}{2\sum_i f_i(1-f_i)}\sum_{i,j}[\mathbf{e}\cdot(\mathbf{x}_i-\mathbf{x}_j)]^2\tilde{\varphi}_{ij}(s) , \tag{2.17}$$

where $\mathbf{e}$ is an arbitrary unit vector and $\tilde{\varphi}_{ij}(s)$ is the Laplace transform of the correlator

$$\varphi_{ij}(t)=\langle\delta n_j(t),\delta n_i\rangle=\langle\delta n_j,e^{-At}\delta n_i\rangle . \tag{2.18}$$

The expression for D is immediate from the Kubo formula for the conductivity σ,[2,11] together with the Einstein relation between σ and D. (Usually, in the frame of the Kubo theory s approaches the origin along the imaginary axis $s=i\omega,\omega\to 0$.) Another common definition of D is given by the linear asymptotic behavior of the second moment of the particle-density distribution. In the terms of the Laplace transforms this is precisely the content of Eq. (2.17) since $\varphi_{ij}(t)$ can be interpreted as the excess carrier distribution in $\mathbf{x}_j$ at time t for a pointlike initial condition in $\mathbf{x}_i$

$$\begin{aligned}\varphi_{ij}(t)&=(\delta n_j,e^{-\mathcal{A}t}P^0\delta n_i)\\&\equiv(\delta n_j,P(t))=\overline{\delta n_j(t)}\end{aligned} \tag{2.19}$$

with

$$P(0)=\delta n_i P^0 , \tag{2.20}$$

so that

$$\varphi_{ij}(0)=f_i(1-f_i)\delta_{ij} . \tag{2.21}$$

It is obvious now that the problem we have to handle is the calculation of

$$\tilde{\varphi}_{ij}(s)=\left\langle\delta n_j,\frac{1}{s+A}\delta n_i\right\rangle . \tag{2.22}$$

The action of A is conveniently described in the orthonormal basis (in $\langle\ ,\ \rangle$) defined as follows.[5] Let

$$\eta_i=\frac{1}{a_i}\delta n_i,\quad a_i=\sqrt{f_i(1-f_i)} . \tag{2.23}$$

For each subset I of the lattice define

$$|\eta_I\rangle=\prod_{i\in I}\eta_i . \tag{2.24}$$

The set $\{|\eta_I\rangle\}_I$ is orthonormal and it can be shown as in Ref. 5 that one has

$$\begin{aligned}A|\eta_I\rangle=&\sum_{\substack{k\in I\\ l\notin I}}|\eta_{I\cup\{l\}\setminus\{k\}}\rangle\Gamma_{k,l}\\&+\sum_{\substack{k\in I\\ l\notin I}}|\eta_{I\cup\{l\}}\rangle\frac{1}{a_k}\Omega_{kl}+\sum_{\substack{k\in I\\ l\in I}}|\eta_{I\setminus\{l\}}\rangle\frac{1}{a_k}\Omega_{kl}\\&+|\eta_I\rangle\left[\sum_{\substack{k\in I,\\ l\notin I}}\frac{1}{a_k^2}W_{kl}+\tfrac{1}{2}\sum_{\substack{k\in I\\ l\in I}}W_{kl}\frac{(f_k-f_l)^2}{a_k^2a_l^2}\right] ,\end{aligned} \tag{2.25}$$

where

$$\Gamma_{kl}=\frac{1}{a_k}M_{kl}\frac{1}{a_l}=\frac{1}{a_k}\left[\delta_{kl}\sum_{l'}W_{kl'}-W_{kl}\right]\frac{1}{a_l}=\Gamma_{lk} , \tag{2.26a}$$

$$\Omega_{kl}=W_{kl}\frac{f_k-f_l}{a_ka_l}=-\Omega_{lk} . \tag{2.26b}$$

Particular cases of (2.25) are

$$A|\eta_j\rangle=\sum_i|\eta_i\rangle\Gamma_{ij}+\sum_i|\eta_i\eta_j\rangle\Omega_{ji}\frac{1}{a_j} \tag{2.27a}$$

[in which, up to some renotations, one can identify the operators appearing in the right-hand side of Eq. (2.9)] and

$$A|\eta_i\eta_j\rangle=\sum_{l\neq i,j}(|\eta_j\eta_l\rangle\Gamma_{il}+|\eta_i\eta_l\rangle\Gamma_{jl})+|\eta_i\rangle\frac{1}{a_i}\Omega_{ij}+|\eta_j\rangle\frac{1}{a_j}\Omega_{ji}$$

$$+\sum_{l\neq i,j}|\eta_i\eta_j\eta_l\rangle\left[\frac{1}{a_i}\Omega_{il}+\frac{1}{a_j}\Omega_{jl}\right]+|\eta_i\eta_j\rangle\left[\sum_{l\neq i,j}\left[\frac{1}{a_i^2}W_{il}+\frac{1}{a_j^2}W_{jl}\right]+W_{ij}\frac{(f_i-f_j)^2}{a_i^2a_j^2}\right]. \tag{2.27b}$$

Several facts are worth discussing here.

(a) In uniform lattices $\Omega_{ij}=0$ and the subspace S_1 of "one-particle" states, spanned by $\{|\eta_i\rangle\}_i$, is invariant for A. Therefore

$$\tilde{\varphi}_{ij}(s)=a_i\left\langle\eta_i,\frac{1}{s+A}\eta_j\right\rangle a_j=a_i\left[\frac{1}{s+\Gamma}\right]_{ij}a_j \tag{2.28}$$

and the "many-body" problem, defined by A in 2^N dimensions becomes a "one-body" problem defined by Γ in the N-dimensional space S_1 (N being the number of sites).

In nonuniform lattices Eq. (2.28) is no longer valid. One defines instead,[5] in analogy with Eq. (2.28), the matrix $\mathcal{H}_{ij}(s)$ by

$$\left\langle\eta_i,\frac{1}{s+A}\eta_j\right\rangle=\left[\frac{1}{s+\mathcal{H}(s)}\right]_{ij}. \tag{2.29}$$

The MF approximation amounts to $\mathcal{H}(s)\simeq\Gamma$, or equivalently

$$Q_1\frac{1}{s+A}Q_1\simeq Q_1\frac{1}{s+Q_1AQ_1}Q_1\ , \tag{2.30}$$

where Q_1 is the projector on S_1.

(b) The $s\to 0$ limit is strongly connected to the spectral behavior of $\mathcal{H}(s)$ near the origin. The following properties are immediate: (i) $\mathcal{H}(s)$ is a Hermitic positive definite matrix; (ii) zero is a nondegenerate eigenvalue of $\mathcal{H}(s)$, the corresponding eigenvector is s independent and has components proportional to a_i.

The first statement stems from similar properties of A. The second is a consequence of the fact that [see Eqs. (2.26) and (2.27)]

$$A\sum_i|\eta_i\rangle a_i=0\ . \tag{2.31}$$

The uniqueness (in S_1) of this eigenvector of A arises from the following connectivity property of the lattice (assumed throughout the paper): the graph defined on the lattice by the bonds (i,j) with $W_{ij}\neq 0$ is connected.

(c) $D(s)$ defined by Eq. (2.17) is an increasing function of s, for which $D(\infty)$ given by

$$\mathbf{e}D(\infty)\mathbf{e}=\frac{1}{2\sum_i a_i^2}\sum_{i,j}[\mathbf{e}\cdot(\mathbf{x}_i-\mathbf{x}_j)]^2W_{ij} \tag{2.32}$$

is an upper bound. Equation (2.32) is an exact MF results in which correlations play no role.

III. PERIODIC LATTICES

If the hopping takes place on a periodic lattice the sites are located at

$$\mathbf{x}_\alpha(\mathbf{r})=\mathbf{r}+\xi_\alpha,\quad \alpha=1,2,\ldots,L\ , \tag{3.1}$$

where $\mathbf{r}$ runs over the Bravais lattice and ξ_α specifies the position inside the cell. The index i is replaced by $(\mathbf{r},\alpha)$, and all quantities defining the model are assumed to be translation invariant

$$\begin{aligned}&\varepsilon_\alpha(\mathbf{r})=\varepsilon_\alpha,\quad a_\alpha(\mathbf{r})=a_\alpha\ ,\\ &W_{\alpha\alpha'}(\mathbf{r},\mathbf{r}')=W_{\alpha\alpha'}(\mathbf{r}-\mathbf{r}')\ .\end{aligned} \tag{3.2}$$

The reason for considering the periodic lattice is twofold: on the one hand many models of interest are periodic (e.g., see Refs. 6 and 8) and on the other hand some results are independent on the size of the unit cell and therefore are relevant for nonperiodic models too. In order to emphasize such situations we shall sometimes keep the notations of the previous section in which the translation invariance is not explicit [e.g., $\mathbf{x}_i$ for $\mathbf{x}_\alpha(\mathbf{r})$ and so on].

The definition of $D(s)$ is now rewritten as

$$\mathbf{e}D(s)\mathbf{e}=\frac{s^2}{2\sum_\alpha a_\alpha^2}\sum_{\alpha\alpha'}\sum_{\mathbf{r}}[\mathbf{e}\cdot(\mathbf{r}+\xi_\alpha-\xi_{\alpha'})]^2$$
$$\times a_\alpha\left\langle\eta_\alpha(\mathbf{r}),\frac{1}{s+A}\eta_{\alpha'}(0)\right\rangle a_{\alpha'}\ . \tag{3.3}$$

In terms of the (discrete) Fourier transformed quantities the second moment of Eq. (3.3) becomes the second derivative at the origin, with respect to the reciprocal-space variable $\mathbf{q}$

$$\mathbf{e}D(s)\mathbf{e}=-\frac{s^2}{2}\left[\mathbf{e}\cdot\frac{\partial}{\partial\mathbf{q}}\right]^2\left\langle\eta(\mathbf{q}),\frac{1}{s+A}\eta(\mathbf{q})\right\rangle\Big|_{\mathbf{q}=0}, \tag{3.4}$$

where

$$|\eta(\mathbf{q})\rangle=\frac{1}{\sqrt{N}}\frac{1}{\left[\sum_\alpha a_\alpha^2\right]^{1/2}}\sum_{\alpha,\mathbf{r}}|\eta_\alpha(\mathbf{r})\rangle a_\alpha e^{i\mathbf{q}\cdot(\mathbf{r}+\xi_\alpha)} \tag{3.5}$$

and N is the number of cells used in the Born–von Kármán cyclic boundary condition.

In terms of the matrix $\mathcal{H}$, and using the canonical scalar product with respect to the α indices, we have

$$\left\langle\eta(\mathbf{q}),\frac{1}{s+A}\eta(\mathbf{q})\right\rangle=\left[u(\mathbf{q}),\frac{1}{s+\mathcal{H}(\mathbf{q},s)}u(\mathbf{q})\right] \tag{3.6}$$

with

$$u_\alpha(\mathbf{q})=\frac{a_\alpha}{\left[\sum_\alpha a_\alpha^2\right]^{1/2}}e^{i\mathbf{q}\cdot\xi_\alpha},\quad \alpha=1,\ldots,L \tag{3.7}$$

and

$$\mathcal{K}_{\alpha\alpha'}(\mathbf{q},s)=\sum_{\mathbf{r}}\mathcal{K}_{\alpha\alpha'}(\mathbf{r},s)e^{-i\mathbf{q}\cdot\mathbf{r}},\quad \alpha,\alpha'=1,\ldots,L\ . \tag{3.8}$$

Therefore the small-$\mathbf{q}$, small-s behavior of the diagonal matrix element of $(s+A)^{-1}$ depends on the "band structure" of $\mathcal{K}(\mathbf{q},s)$. The properties of $\mathcal{K}(s)$ listed in the previous section are translated here as follows: for every $\mathbf{q}$ and s the spectrum of $\mathcal{K}(\mathbf{q},s)$ is strictly positive, except for the lowest branch, which touches the origin at $\mathbf{q}=0$, and has a parabolic behavior around this point. More precisely, we have the spectral decomposition of $\mathcal{K}(\mathbf{q},s)$

$$\mathcal{K}(\mathbf{q},s)=\sum_{\lambda=1}^{L}k_\lambda(\mathbf{q},s)|v_\lambda(\mathbf{q},s))(v_\lambda(\mathbf{q},s)| \tag{3.9}$$

with the following behavior near $\mathbf{q}=0$:

$$\begin{aligned}&k_\lambda(\mathbf{q},s)\sim k_\lambda(0,s)>0,\quad \lambda>1\ ,\\ &k_1(\mathbf{q},s)\sim \mathbf{q}\mathcal{D}(s)\mathbf{q}\ ,\\ &|v_1(\mathbf{q},s))\sim\left[\sum_\alpha a_\alpha^2\right]^{-1/2}|a)\ ,\end{aligned} \tag{3.10}$$

where $\mathcal{D}(s)$ is a positive-definite tensor and $|a)$ the vector with components equal to a_α. Around $s=0$, $\mathbf{q}=0$ the leading (singular) term of Eq. (3.6) stems from the lowest branch of the spectrum ($\lambda=1$)

$$\begin{aligned}\left[u(\mathbf{q}),\frac{1}{s+\mathcal{K}(\mathbf{q},s)}u(\mathbf{q})\right]&=\sum_{\lambda=1}^{L}\frac{1}{s+k_\lambda(\mathbf{q},s)}|(u(\mathbf{q}),v_\lambda(\mathbf{q},s))|^2\\ &\sim\frac{1}{s+\mathbf{q}\mathcal{D}(0)\mathbf{q}}\end{aligned} \tag{3.11}$$

and is the only one that survives in the $s\to 0$ limit of Eq. (3.4) giving

$$D=\mathcal{D}(0)\ . \tag{3.12}$$

Two aspects are essential here.

(a) Equation (3.11) shows that the excess carrier distribution has the expected diffusive behavior around $s=0$, $\mathbf{q}=0$. Other details of the distribution are irrelevant. The whole procedure, involving the second derivative followed by the limits $\mathbf{q}\to 0$ and $s\to 0$, is equivalent to the statement that the diffusion constant is equal to the tensor appearing in the right-hand side of Eq. (3.11).

(b) In the proof of Eq. (3.11) the only important feature of $|u(\mathbf{q}))$ is the fact that its leading term in the $q\to 0$ limit coincides with that of $|v_1(\mathbf{q},s))$. Therefore we may replace $|u(\mathbf{q}))$ *by any other vector having the same behavior* without changing the result. Nevertheless the approximation procedure is affected by the change and one can use this degree of freedom in order to optimize the scheme, in a sense that will be clarified below.

The diagonal matrix element of the resolvent of A from Eq. (3.4) can be represented in a continued-fraction expansion using the recursion method.[17] To this end, starting with $|\eta_1\rangle=|\eta(\mathbf{q})\rangle$, by repeated applications of A one generates an orthonormal basis $|\eta_2\rangle$, $|\eta_3\rangle,\ldots$, in which A takes the form of a Jacobi matrix

$$\begin{aligned}A|\eta_1(\mathbf{q})\rangle&=a_1(\mathbf{q})|\eta_1(\mathbf{q})\rangle+b_1(\mathbf{q})|\eta_2(\mathbf{q})\rangle\ ,\\ A|\eta_2(\mathbf{q})\rangle&=b_1^*(\mathbf{q})|\eta_1(\mathbf{q})\rangle+a_2(\mathbf{q})|\eta_2(q)\rangle\\ &\quad+b_2(\mathbf{q})|\eta_3(\mathbf{q})\rangle\ ,\end{aligned} \tag{3.13}$$

and so on. Then one has

$$\left\langle\eta_1(\mathbf{q}),\frac{1}{s+A}\eta_1(\mathbf{q})\right\rangle=\cfrac{1}{s+a_1(\mathbf{q})-\cfrac{|b_1(\mathbf{q})|^2}{s+a_2(\mathbf{q})-\cfrac{|b_2(q)|^2}{\cdots}}}\ . \tag{3.14}$$

The behavior discussed above [Eq. (3.11)] is ensured by the fact that a_1 is quadratic in $\mathbf{q}$ and b_1 is linear. All the other terms in the expansion may be taken at their $\mathbf{q}=0$, $s=0$ value.

The complete calculation of the continued fraction is usually impossible. Truncating it at different levels p by putting $b_p=0$ leads, due to the positivity of A, to a decreasing sequence of upper bounds for the true value of the diffusion constant: $D^{(1)}\geq D^{(2)}\geq\cdots\geq D^{(p)}\geq\cdots\geq D$.

Using [see Eq. (3.5)]

$$|\eta_1(\mathbf{q})\rangle=\left[\sum_i a_i^2\right]^{-1/2}\sum_i|\eta_i\rangle a_i e^{i\mathbf{q}\cdot\mathbf{x}_i} \tag{3.15}$$

together with Eqs. (2.26) and (2.27a), one immediately obtains

$$a_1(\mathbf{q})=\frac{1}{\sum_i a_i^2}\sum_{i,j}M_{ij}e^{-i\mathbf{q}\cdot(\mathbf{x}_i-\mathbf{x}_j)}\ . \tag{3.16}$$

The leading term of $a_1(\mathbf{q})$ when $\mathbf{q}\to 0$ along the $\mathbf{e}$ direction ($\mathbf{q}=q\mathbf{e}$, $q\to 0$) is

$$\begin{aligned}a_1(\mathbf{q})&\sim q^2\frac{1}{2\sum_i a_i^2}\sum_{i,j}[\mathbf{e}\cdot(\mathbf{x}_i-\mathbf{x}_j)]^2W_{ij}\\ &=q^2\mathbf{e}D(\infty)\mathbf{e}\ .\end{aligned} \tag{3.17}$$

Therefore the first term in this sequence of upper bounds for D is $D^{(1)}=D(\infty)$. This is not a very good starting point for our approximation scheme, being higher than the MF value D^{MF}. It is very important to optimize the first step of the recursion procedure because the repeated applications of A, though straightforward in principle, lead to tedious calculations.

In order to improve $D^{(1)}$ let us consider the spectral decomposition of $\Gamma(\mathbf{q})$,[3]

$$\Gamma(\mathbf{q})=\sum_{\lambda=1}^{L}\gamma_\lambda(\mathbf{q})|u_\lambda(\mathbf{q}))(u_\lambda(\mathbf{q})| \tag{3.18}$$

which has strong similarities with that of $\mathcal{K}(\mathbf{q},s)$ in what concerns the $\mathbf{q}\to 0$ behavior,

$$\gamma_\lambda(\mathbf{q})\sim\gamma_\lambda(0)>0,\quad \lambda\neq 1$$
$$\gamma_1(\mathbf{q})\sim \mathbf{q}D^{\mathrm{MF}}\mathbf{q}\ , \tag{3.19}$$
$$|u_1(\mathbf{q})\rangle\sim\left[\sum_\alpha a_\alpha^2\right]^{-1/2}|a\rangle\ .$$

It is clear that $|u_1(\mathbf{q})\rangle$ is a valid replacement for $|u(\mathbf{q})\rangle$ since both vectors have the same leading term. Changing in the expression of $\mathbf{e}D(s)\mathbf{e}\,|u(\mathbf{q})\rangle$ by $|u_1(\mathbf{q})\rangle$ does not affect the value of D, but gives the lowest possible value of the first approximant $D^{(1)}$, because we approach $\mathbf{q}=0$ along the lowest spectral branch of $\Gamma(\mathbf{q})$. Corresponding to this change the first term in the orthonormal basis of the recursion method is now

$$|\eta_1(\mathbf{q})\rangle=\frac{1}{\sqrt{N}}\sum_{\alpha,\mathbf{r}}|\eta_\alpha(\mathbf{r})\rangle e^{i\mathbf{q}\cdot\mathbf{r}}u_{1\alpha}(\mathbf{q}) \tag{3.20}$$

and leads to

$$A|\eta_1(\mathbf{q})\rangle=\gamma_1(\mathbf{q})|\eta_1(\mathbf{q})\rangle+\frac{1}{2\sqrt{N}}\sum_{\substack{\alpha,\alpha',\\ \mathbf{r},\mathbf{r}'}}|\eta_\alpha(\mathbf{r})\eta_{\alpha'}(\mathbf{r}')\rangle\Omega_{\alpha\alpha'}(\mathbf{r}-\mathbf{r}')\left[e^{i\mathbf{q}\cdot\mathbf{r}}\frac{u_{1\alpha}(\mathbf{q})}{a_\alpha}-e^{i\mathbf{q}\cdot\mathbf{r}'}\frac{u_{1\alpha'}(\mathbf{q})}{a_{\alpha'}}\right]\ . \tag{3.21}$$

In the first term one identifies

$$a_1(\mathbf{q})=\gamma_1(\mathbf{q}) \tag{3.22}$$

and therefore [see Eq. (3.19)] we have now

$$D^{(1)}=D^{\mathrm{MF}}\ . \tag{3.23}$$

The second term is orthogonal on the first and consequently it is equal to $b_1(\mathbf{q})|\eta_2(\mathbf{q})\rangle$. We need only its leading (first) order in $\mathbf{q}$,

$$b_1(\mathbf{q})|\eta_2(\mathbf{q})\rangle\sim\frac{iq}{2\left[\sum_i a_i^2\right]^{1/2}}\sum_{i,j}|\eta_i\eta_j\rangle\Omega_{ij}\mathbf{e}\cdot(\mathbf{y}_i-\mathbf{y}_j)\ , \tag{3.24}$$

where we used the expansion of $|u_1(\mathbf{q})\rangle$

$$u_{1,\alpha}(\mathbf{q})\sim\frac{a_\alpha}{\left[\sum_\alpha a_\alpha^2\right]^{1/2}}[1+i\mathbf{q}\cdot\mathbf{c}_\alpha] \tag{3.25}$$

up to the first order in $\mathbf{q}$, together with the definition of $\mathbf{y}_i$

$$\mathbf{y}_\alpha(\mathbf{r})=\mathbf{r}+\mathbf{c}_\alpha\ . \tag{3.26}$$

The only relevant information contained in $|u_1(\mathbf{q})\rangle$ is the set of vectors $\mathbf{c}_\alpha$, defined by Eq. (3.25) or equivalently the set $\mathbf{y}_\alpha(\mathbf{r})$ of Eq. (3.26). These are MF theoretic quantities, i.e., they are entirely defined by $\Gamma(\mathbf{q})$, and once they are known $|\eta_2(0)\rangle$ is known, and the recursion method proceeds with iterated applications of A upon it.

Although the definition of $\{\mathbf{y}_i\}_i$ seems to rely on the translational invariance of the lattice, a closer look at the problem shows that this is not so. Indeed, expanding

$$\Gamma(\mathbf{q})|u_1(\mathbf{q})\rangle=\gamma_1(\mathbf{q})|u_1(\mathbf{q})\rangle \tag{3.27}$$

in powers of $\mathbf{q}$ up to first order, and using Eqs. (3.19) and (3.25), we get

$$\sum_{\alpha',\mathbf{r}'}M_{\alpha\alpha'}(\mathbf{r}-\mathbf{r}')\mathbf{y}_{\alpha'}(\mathbf{r}')=0 \tag{3.28}$$

or, equivalently

$$\sum_j M_{ij}\mathbf{y}_j=\sum_j W_{ij}(\mathbf{y}_i-\mathbf{y}_j)=0\ . \tag{3.29}$$

Therefore $\{\mathbf{y}_i\}_i$ is the solution of $M\mathbf{y}=0$, having the asymptotic behavior [compare Eqs. (3.1) and (3.26)]

$$\mathbf{y}_i\underset{\mathbf{x}_i\to\infty}{\sim}\mathbf{x}_i\ . \tag{3.30}$$

If we interpret the operator M as a discretization of the Laplacian, this is the counterpart of the linear solution of the Laplace equation. Local inhomogeneities lead to local fluctuations around the purely linear solution which remains only asymptotically valid. In the stationary flow theories the departures of $\mathbf{y}_i$ from $\mathbf{x}_i$ are related to the well-known local changes in the chemical potential.[1,18]

Fortunately the periodic lattices are not the only ones for which asymptotic linear solution is available. Another interesting class of models with this property consists of linear chains, under quite reasonable assumptions, as we shall see in the next section. On physical grounds one may assume the existence of such a solution for nonperiodic systems, provided they are macroscopically homogeneous.

All the quantities needed for the continued-fraction expansion may be expressed in terms of the asymptotic linear solution. For instance,

$$a_1(\mathbf{q})=\gamma_1(\mathbf{q})=(u_1(\mathbf{q}),\Gamma(\mathbf{q})u_1(\mathbf{q}))\sim\frac{q^2}{2N\sum_\alpha a_\alpha^2}\sum_{\substack{\alpha,\alpha'\\ \mathbf{r},\mathbf{r}'}}W_{\alpha\alpha'}(\mathbf{r}-\mathbf{r}')[\mathbf{e}\cdot(\mathbf{y}_\alpha(\mathbf{r})-\mathbf{y}_{\alpha'}(\mathbf{r}'))]^2\ , \tag{3.31}$$

wherefrom

$$\mathbf{e}D^{\mathrm{MF}}\mathbf{e}=\frac{1}{2\sum_i a_i^2}\sum_{i,j}[\mathbf{e}\cdot(\mathbf{y}_i-\mathbf{y}_j)]^2W_{ij}\ . \tag{3.32}$$

Also, from Eq. (3.24)

$$|b_1(\mathbf{q})|^2\sim q^2\frac{1}{2\sum_i a_i^2}\sum_{i,j}[\mathbf{e}\cdot(\mathbf{y}_i-\mathbf{y}_j)]^2\Omega_{ij}^2 \tag{3.33}$$

and $|\eta_2(0)\rangle$ is (up to normalization)

$$|\eta_2(0)\rangle=\sum_{i,j}|\eta_i\eta_j\rangle\Omega_{ij}e^{\mathbf{e}\cdot(\mathbf{y}_i-\mathbf{y}_j)}\ . \tag{3.34}$$

IV. SOME APPLICATIONS

A. Exact solubility at the MF level

It is natural to ask whether there are models for which the continued fraction is finite due to the vanishing of some b in Eq. (3.14). Of course such models are exactly soluble.

All the lattices with identical site energies have $b_1=0$ because $\Omega_{ij}=0$ and therefore the MF theory becomes exact. Nevertheless the class of models for which $D=D^{\mathrm{MF}}$ (at least in some direction) is larger as can be seen from the examination of Eq. (3.33). It is sufficient to have $\mathbf{e}\cdot\mathbf{y}_i=\mathbf{e}\cdot\mathbf{y}_j$ whenever $\Omega_{ij}\neq 0$. This situation can be illustrated on the model of Ref. 8. In this model the hopping sites are arranged in a tetragonal lattice, and the diffusion constant along the c axis was found in Ref. 8 to be equal to the MF value in a certain domain of the parameters. In fact the MF theory is exact for any values of the hopping rates, because $\Omega_{ij}\neq 0$ only for i and j on the same (001) plane, on which $\mathbf{e}\cdot\mathbf{y}_i=\mathbf{e}\cdot\mathbf{x}_i=\text{const}$ for $\mathbf{e}=(001)$ due to the reflection symmetry with respect to the plane.

B. The Richards model

Consider now the model of Ref. 6. It consists of a linear chain of equidistant alternating sites of two types A and B. The hopping takes place only between nearest neighbors. Only one value of W is present,

$$W=\mathcal{W}_{AB}f_B(1-f_A)=\mathcal{W}_{BA}f_A(1-f_B)\ . \tag{4.1}$$

Straightforward computations give (d is the spacing between the sites)

$$\begin{aligned}
&D^{\mathrm{MF}}=\frac{2Wd^2}{a_A^2+a_B^2}\ ,\\
&|b_1(q)|^2\sim q^2\frac{2Wd^2}{(a_A^2+a_B^2)a_A^2a_B^2}(f_A-f_B)^2\ ,\\
&a_2=W\frac{(f_A-f_B)^2+a_A^2+a_B^2}{a_A^2a_B^2}\ ,\\
&b_2=0\ .
\end{aligned} \tag{4.2}$$

Due to the vanishing of b_2 the first correction to D^{MF} gives the exact result

$$D=D^{(2)}=D^{\mathrm{MF}}\frac{a_A^2+a_B^2}{(f_A-f_B)^2+a_A^2+a_B^2}\ , \tag{4.3}$$

which was first obtained by Richards[6] using a stationary-flow formalism. The latter is expected to be equivalent to the transitory-regime approach used here, as suggested by the above results, but the proof is not immediate (see Ref. 16). General conditions under which $b_2=0$ can be written down but are difficult to interpret and it seems that they are not met by other simple models.

C. The linear chain

Assuming for simplicity the sites are equidistant with distance d, and denoting $W_{i,i+1}=W_{i+1,i}=W_i$, Eq. (3.29) takes the form

$$W_{i-1}(y_i-y_{i-1})=W_i(y_{i+1}-y_i)\ , \tag{4.4}$$

wherefrom

$$y_i=\alpha\sum_{k<i}\frac{1}{W_k}\ . \tag{4.5}$$

The correct asymptotic behavior $y_i\sim id$ for $i\to\infty$ [Eq. (3.30)] is ensured by taking $\alpha=d\langle W^{-1}\rangle^{-1}$ where

$$\langle W^{-1}\rangle=\lim_{i\to\infty}\frac{1}{i}\sum_{k<i}\frac{1}{W_k}\ . \tag{4.6}$$

The existence of the average for the inverse hopping rates given by the above limit is not a too restrictive hypothesis and will be assumed to hold in what follows.

Using Eq. (3.32) one obtains

$$D^{\mathrm{MF}}=\frac{d^2}{\langle a^2\rangle}\langle W^{-1}\rangle^{-1}\ , \tag{4.7}$$

with the average $\langle\ \rangle$ defined as in Eq. (4.6) for all quantities. Equation (4.7) is a well-known result.[4] The first correlation corrections involve more complicated averages which depend on the relative frequency of the apparition on the chain of different types of atoms and sequences of atoms. For the case of the Richards model we recover the results of Eq. (4.3). Another simple case is that of a binary random chain for which

$$D^{(2)}=D^{\mathrm{MF}}\left[1-\frac{2p_Ap_B(f_A-f_B)^2\langle W^{-1}\rangle^{-1}}{p_Aa_B^2W_{AA}+p_Ba_A^2W_{BB}+[p_Aa_A^2+p_Ba_B^2+(f_A-f_B)^2]W_{AB}}\right]\ , \tag{4.8}$$

where p_A and p_B are the concentrations of the atoms A and B, respectively. Therefore,

$$\langle W^{-1}\rangle = p_A^2 \frac{1}{W_{AA}} + p_B^2 \frac{1}{W_{BB}} + 2p_A p_B \frac{1}{W_{AB}} ,$$
$$\langle a^2\rangle = p_A a_A^2 + p_B a_B^2 . \tag{4.9}$$

V. VARIATIONAL FORMULATION OF THE HOPPING DIFFUSION PROBLEM

The frequent appearance of upper bounds for the diffusion constant in the hopping model is connected to the fact that the problem can be formulated variationally. The variational principle is quite simple and reads

$$\left\langle \eta_1(\mathbf{q}), \frac{1}{s+A}\eta_1(\mathbf{q})\right\rangle = \sup_x [2\ \mathrm{Re}\langle x,\eta_1(q)\rangle - \langle x,(s+A)x\rangle] \tag{5.1}$$

where the supremum is taken over all the states in the Hilbert space in which the operator A acts.

We can rewrite the variational problem in a way in which the functional does not depend on the normalization of $|x\rangle$. To this end we change $|x\rangle$ into $\lambda|x\rangle$ in Eq. (5.1) and take the supremum with respect to $\lambda\in\mathbb{C}$.

This yields

$$\left\langle \eta_1(\mathbf{q}), \frac{1}{s+A}\eta_1(\mathbf{q})\right\rangle = \frac{1}{\inf_x \dfrac{\langle x,(s+A)x\rangle}{|\langle x,\eta_1(\mathbf{q})\rangle|^2}} . \tag{5.2}$$

Now one may normalize $|x\rangle$ so that $\langle x,\eta_1(q)\rangle = 1$ [of course one may also have states with $\langle x,\eta_1(\mathbf{q})\rangle = 0$ but these cannot be solutions of our variational problem]. Therefore one may write

$$|x\rangle = |\eta_1(q)\rangle + |x_1\rangle \tag{5.3}$$

with $|x_1\rangle$ orthogonal on $|\eta_1(\mathbf{q})\rangle$. Equation (5.2) becomes

$$\left\langle \eta_1(\mathbf{q}), \frac{1}{s+A}\eta_1(\mathbf{q})\right\rangle = \frac{1}{s+a_1(\mathbf{q}) + \inf\limits_{x_1\perp\eta_1(\mathbf{q})} [2\ \mathrm{Re}\langle \eta_1(\mathbf{q}),(s+A)x_1\rangle + \langle x_1,(s+A)x_1\rangle]} . \tag{5.4}$$

Using again $\mathbf{q}=q\mathbf{e}$ and

$$A|\eta_1(\mathbf{q})\rangle = a_1(\mathbf{q})|\eta_1(\mathbf{q})\rangle + q|\eta_2(\mathbf{q})\rangle \tag{5.5}$$

as in Eq. (3.13) [but this time with $|\eta_2(\mathbf{q})\rangle$ not normalized] we obtain, after changing our variational parameter $|x_1\rangle$ into $-q|x_1\rangle$,

$$\left\langle \eta_1(\mathbf{q}), \frac{1}{s+A}\eta_1(\mathbf{q})\right\rangle = \frac{1}{s+a_1(\mathbf{q}) - q^2 \sup\limits_{x_1\perp\eta_1(\mathbf{q})} [2\ \mathrm{Re}\langle \eta_2(q),x_1\rangle - \langle x_1,(s+A)x_1\rangle]} , \tag{5.6}$$

wherefrom

$$\mathbf{e}D\mathbf{e} = \mathbf{e}D^{\mathrm{MF}}\mathbf{e} - \sup_{x_1\perp\eta_1(\mathbf{q})} [2\mathrm{Re}\langle \eta_2(0),x_1\rangle - \langle x_1, Ax_1\rangle] . \tag{5.7}$$

In Eq. (5.6) we got another variational problem similar to the initial one. By repeating the steps which led us from Eq. (5.1) to Eq. (5.6) we gradually recover the continued fraction described in Sec. III. Terminating it by zero at a certain level amounts to taking the corresponding trial vector equal to zero (as, for instance, the MF approximation corresponds to choosing $|x_1\rangle = 0$). Of course other choices may be made, all leading to lower bounds for the true supremum, i.e., upper bounds for the diffusion constant.

In what follows we take advantage of the flexibility of the variational principle in order to improve the approximation scheme. Denoting as in Sec. II by S_1 the space spanned by the one-particle states $|\eta_i\rangle$ and introducing the subspace S_2 spanned by the "W-bond" states $|\eta_i\eta_j\rangle$ where (i,j) are pairs of indices with $W_{ij}\neq 0$, we observe that S_1 and S_2 are orthogonal and $|\eta_1(\mathbf{q})\rangle\in S_1$ and $|\eta_2(\mathbf{q})\rangle\in S_2$.

Corresponding to the splitting of the whole space in the direct sum of S_1, S_2, and their orthogonal S_3 (in this order), we have the 3×3 block matrix form of A,

$$A = \begin{bmatrix} \Gamma & \Omega^\dagger & 0 \\ \Omega & T & X^\dagger \\ 0 & X & V \end{bmatrix} , \tag{5.8}$$

where the structure is immediate from Eqs. (2.25)–(2.27) and the blocks may be identified there. If in Eq. (5.2) we restrict the search of the infimum to S_1, we obtain

$$\inf_x \frac{\langle x,(s+A)x\rangle}{|\langle x,\eta_1(\mathbf{q})\rangle|^2} \simeq \inf_{x_1\in S_1} \frac{\langle x_1,(s+\Gamma)x_1\rangle}{|\langle x_1,\eta_1(\mathbf{q})\rangle|^2} , \tag{5.9}$$

which is the MF approximation. By extending the search to $S_1\oplus S_2$ we get a better result,

$$\inf_x \frac{\langle x,(s+A)x\rangle}{|\langle x,\eta_1(\mathbf{q})\rangle|^2} \simeq \inf_{x_1\in S_1}\inf_{x_2\in S_2} \frac{\langle x_1,(s+\Gamma)x_1\rangle + 2\ \mathrm{Re}\langle x_2,\Omega x_1\rangle + \langle x_2,(s+T)x_2\rangle}{|\langle x_1,\eta_1(\mathbf{q})\rangle|^2} . \tag{5.10}$$

The infimum in S_2 is achieved for

$$|x_2\rangle = -\frac{1}{s+T}\Omega|x_1\rangle \tag{5.11}$$

and we are left with finding

$$\inf_{x_1\in S_1} \frac{\left\langle x_1, \left[s+\Gamma-\Omega^\dagger\frac{1}{s+T}\Omega\right]x_1\right\rangle}{|\langle x_1,\eta_1(\mathbf{q})\rangle|^2} . \tag{5.12}$$

This is quite similar to the MF problem Eq. (5.9) with the "effective" Γ (only its $s=0$ value enters the calculation):

$$\tilde{\Gamma}=\Gamma-\Omega^\dagger\frac{1}{T}\Omega . \tag{5.13}$$

Inverting T is in general a difficult problem. In periodic lattices, if the number of "bonds" is finite (up to translations), the Fourier transform reduces it to the inversion of a finite matrix, but the result depends strongly on the particular geometry and on the size of the elementary cell.

Therefore it is remarkable that *in lattices in which the W bonds form no triangles* the inversion of T becomes trivial. Indeed, the examination of Eq. (2.27b) shows that in such lattices T is a diagonal operator in S_2,

$$T_{ij;kl}=\delta_{(ij)(kl)}T_{ij} , \tag{5.14a}$$

$$T_{ij}=W_{ij}\frac{(f_i-f_j)^2}{a_i^2a_j^2}+\sum_{l\neq i,j}\left[\frac{1}{a_i^2}W_{il}+\frac{1}{a_j^2}W_{jl}\right] . \tag{5.14b}$$

The improved MF problem defined by $\tilde{\Gamma}$ corresponds to renormalized transition rates $\tilde{W}_{ij}$ given by

$$\frac{1}{\tilde{W}_{ij}}=\frac{1}{W_{ij}}+\frac{(f_i-f_j)^2}{\sum_{l\neq i,j}(a_j^2W_{il}+a_i^2W_{jl})} . \tag{5.15}$$

All the advantages of the MF theory are present, including the equivalence *à la* Miller and Abrahams[1] with a resistor network problem. Equation (5.15) says how to modify the resistors to include the correlation effects. The upper bound $\tilde{D}^{\mathrm{MF}}$ obtained in this way is not only lower than the usual MF but also lower than the first correction to it $D^{(2)}$. This happens because $D^{(2)}$ corresponds to taking $b_2=0$ which amounts to searching the infimum of Eq. (5.2) in the subspace generated by $|\eta_1\rangle$ and $|\eta_2\rangle$, while here we minimized on the whole $S_1\oplus S_2$ subspace.

Let us make the following remarks.

(a) If triangles are present in the lattice, T is no longer diagonal. Denoting its diagonal part by T^Δ it is clear that

$$|x_2\rangle = -\frac{1}{s+T^\Delta}|x_1\rangle \tag{5.16}$$

is no longer the optimal choice of the S_2 component. Nevertheless Eq. (5.16) gives a trial vector for our variational problem. Whether this guess is good or not depends on how large the off-diagonal part of T is. The result for the "improved" transition rates in this case reads

$$\tilde{W}_{ij}=W_{ij}-\frac{W_{ij}^2}{a_i^2a_j^2}\frac{(f_i-f_j)^2}{T_{ij}} + \sum_{l\neq i,j}\frac{W_{il}W_{lj}W_{ij}}{a_i^2a_j^2a_l^2}\left[\frac{(f_i-f_l)(f_l-f_j)}{T_{il}T_{lj}}+\frac{(f_i-f_j)(f_l-f_j)}{T_{ij}T_{lj}}+\frac{(f_i-f_l)(f_i-f_j)}{T_{il}T_{ij}}\right] . \tag{5.17}$$

The last sum in Eq. (5.17) is the contribution of the triangles. If these are missing we are left with the first two terms, which give Eq. (5.15).

(b) It is interesting to note that results which turn out to be equivalent to this improved MF theory were obtained earlier, using the stationary-flow approach, in two particular cases: in ordered two-sublattice structures[6,9] with hops only between the sublattices, and in one-dimensional chains.[15] In both cases the lattices had no triangles.

(c) Corrections to $\tilde{D}^{\mathrm{MF}}$ can be also obtained using a continued-fraction expansion, as in Sec. III. Truncations of the fraction also lead to a decreasing sequence of upper bounds for D: $\tilde{D}^1=\tilde{D}^{\mathrm{MF}}\geq\tilde{D}^{(2)}\geq\cdots\geq D$. As explained above we have $\tilde{D}^{(1)}\leq D^{(2)}$.

In order to prove these statements it is enough to notice that taking the infimum in Eq. (5.2) with respect to the S_2 component of x eliminates this subspace from the problem, leaving a renormalized A operator acting on $S_1\oplus S_3$,

$$\tilde{A}=\begin{bmatrix} \tilde{\Gamma} & -\Omega^\dagger\frac{1}{T}X^\dagger \\ -X\frac{1}{T}\Omega & V-X\frac{1}{T}X^\dagger \end{bmatrix} . \tag{5.18}$$

Apart from this modification, the whole scheme of Sec. III remains unchanged.

(d) Immediate consequences of the renormalization of the resistors Eq. (5.15) can be seen in the low-temperature behavior of the conductivity. As an illustration we consider the example of the random binary chain with $\varepsilon_A<\varepsilon_B$ at particle concentrations for which the chemical potential lies between ε_A and ε_B (in our convention $\mu=0$ so that $\varepsilon_A<0<\varepsilon_B$). We assume the energy and temperature dependence of the transition rates as in Ref. 18 and

discuss the so-called correlation factor[15] defined as

$$f=\frac{D}{D^{\mathrm{MF}}} . \tag{5.19}$$

In the first correction to the usual MF theory, as given by Eq. (4.8), one immediately obtains the $T\to 0$ limit

$$f^{(2)}=\frac{D^{(2)}}{D^{\mathrm{MF}}}\to\frac{1}{2} , \tag{5.20}$$

indicating a reduction of the diffusion to one-half. In fact, the reduction due to the correlations is much more drastic as shown by the improved MF theory of this section. Indeed, the corresponding correlation factor

$$\tilde{f}=\frac{\tilde{D}^{\mathrm{MF}}}{D^{\mathrm{MF}}}=\frac{\langle \tilde{W}^{-1}\rangle^{-1}}{\langle W^{-1}\rangle^{-1}}\underset{T\to 0}{\sim} e^{-\beta\Delta} \tag{5.21}$$

becomes exponentially small ($\Delta=\min\{|\varepsilon_A|,|\varepsilon_B|\}$ is the shortest distance from the chemical potential to an energy level). The exponential factor of Eq. (5.21) contributes to the increase of the activation energy of the conductivity from its MF value $E=\varepsilon_B-\varepsilon_A$ to the value $E+\Delta$. This effect was found in the literature both analytically and numerically[6,13] for the Richards and the random model, respectively. In fact the improved MF theory predicts the activation energy $E+\Delta$ for any chain in which sequences of four alternating sites have nonzero frequency of appearance. Indeed, according to Eq. (5.15) such sequences give rise to the highest resistances in the chain, increasing as $\exp\{\beta(E+\Delta)\}$ and they control the total conductivity of the sample.

Going to higher corrections one may expect a still higher activation energy, but both numeric and heuristic arguments[13,15] indicate that the lowest conductivity among the binary chains is that of the Richards model, for which $\tilde{D}^{\mathrm{MF}}$ coincides with the exact solution. These are strong reasons to believe that the true activation energy is $E+\Delta$.

VI. SUMMARY AND CONCLUSIONS

We have presented a theory of the diffusion constant D of hopping systems with hard-core interaction based on a formula [Eq. (2.17)] that involves the restriction of the resolvent of the master evolution operator A in the one-particle subspace S_1. We have considered the case of periodic lattices, in which the Laplace-Fourier transform of the diffusion Gaussian emerges as the small-s, small-$\mathbf{q}$ singularity of the excess particle evolution [Eq. (3.11)] and the diffusion constant can be identified in this leading term.

The evaluation of the diagonal matrix element of the resolvent was done by the recursion method. (This is entirely in the spirit of Fedders and Sankey[5] who used Padé approximants for the calculation of each matrix element $\mathcal{H}_{\alpha\alpha'}(\mathbf{q},s)$, with the only difference that we compute successive approximants for D itself). Due to the positivity of A, successive truncations of the continued fraction form a decreasing sequence of upper bounds for D.

It was shown that the essential feature of the state $|u(\mathbf{q})\rangle$, on which the resolvent is calculated, is its $\mathbf{q}\to 0$ limit, identical to that of $|v_1(\mathbf{q},s)\rangle$, the eigenvector of $\mathcal{H}(\mathbf{q},s)$ corresponding to the lowest spectral branch. This behavior (as well as the vanishing of the corresponding eigenvalue) is connected to the conservation of the number of particles in the hopping process. This essential property is preserved in the MF approximation, wherefrom the same $\mathbf{q}\to 0$ limit stems for $|u_1(\mathbf{q})\rangle$ and $\gamma_1(\mathbf{q})$ of Eq. (3.19). Replacing $|u(\mathbf{q})\rangle$ by $|u_1(\mathbf{q})\rangle$ was shown in this way to be legitimate, as far as the exact diffusion constant D is concerned, and an improvement for the approximation procedure, since the first approximant is lowered to the value $D^{(1)}=D^{\mathrm{MF}}$. All the terms of the continued fraction can be written in terms of the asymptotic linear solution of Eq. (3.29). Generalization to nonperiodic systems comes with the realization that the existence of such solutions is not limited to periodic lattices. Nevertheless a word of caution is needed here: extrapolations to nonperiodic systems are done by taking the infinite elementary cell limit on the results, i.e., after the $s\to 0$ ($t\to\infty$) limit, and not the other way around. In other words, no matter how large the elementary cell, the particle has enough time to explore a still larger region and "feels" a periodic environment. The problem of commuting these limits is delicate even in the MF description (see, e.g., Ref. 19) and is beyond the aim of this paper. The variational procedure described in Sec. V was exploited in the sense of improving the above approximation scheme.

By renormalizing the transition rates so as to include correlation effects a modified MF-type description was proposed. It was shown to lead to a better approximation to the diffusion constant than the usual MF theory, while preserving all the advantages of a one-particle description (including the Miller-Abrahams analogy with a network of renormalized resistors). Strong correlation effects were easily pointed out using the improved MF theory in the low-temperature behavior of the conductivity.

Note added in proof. It has been called to our attention that Spohn[20] has also given a variational formulation for the diffusivity problem of lattice gases, in a slightly different setting (uniform periodic lattices and arbitrary short-range interaction). The upper bound obtained in Ref. 20 corresponds to $D(\infty)$.

[1]A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).
[2]L. Bányai and A. Aldea, Fortschr. Phys. **27**, 435 (1979).
[3]L. Bányai and P. Gartner, Physica **102A**, 357 (1980).
[4]S. Alexander, J. Bernasconi, W. R. Schneider, and R. Orbach, Rev. Mod. Phys. **53**, 175 (1981).
[5]P. A. Fedders and O. F. Sankey, Phys. Rev. B **15**, 3580 (1977).
[6]P. M. Richards, Phys. Rev. B **16**, 1393 (1977).
[7]P. A. Fedders, Phys. Rev. B **17**, 2098 (1978).
[8]P. A. Fedders, Phys. Rev. B **18**, 1055 (1978).
[9]P. M. Richards, J. Chem. Phys. **68**, 2125 (1978).

[10]P. M. Richards, in *Fast Ion Transport in Solids,* edited by P. Vashishta *et al.* (North-Holland, Amsterdam, 1979), p. 349.
[11]P. A. Fedders and P. M. Richards, Phys. Rev. B **21**, 377 (1980).
[12]K. Nakazato and K. Kitahara, Prog. Theor. Phys. **64**, 2261 (1980).
[13]K. S. Chase and D. J. Thouless, Phys. Rev. B **39**, 9809 (1989).
[14]W. Dieterich, Solid State Ionics **40/41**, 509 (1990).
[15]R. Pitiş and P. Gartner, Phys. Rev. B **43**, 11 294 (1991).
[16]R. Pitiş and P. Gertner, Rev. Roum. Phys. **36**, 867 (1991).
[17]*The Recursion Method and Its Applications,* edited by D. G. Pettifor and D. L. Weaire (Springer, Berlin, 1985).
[18]V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).
[19]J. P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[20]H. Spohn, in *Proceedings of the 9th International Congress in Mathematical Physics,* edited by B. Simon, A. Truman, and I. M. Davies (Adam Hilger, Bristol, 1989), p. 272.