PHYSICAL REVIEW B **79**, 235206 (2009)

Theory of charge transport in organic crystals: Beyond Holstein's small-polaron model

Frank Ortmann, Friedhelm Bechstedt, and Karsten Hannewald
Institut für Festkörpertheorie und-optik, Friedrich-Schiller-Universität Jena and European Theoretical Spectroscopy Facility, Max-Wien-Platz 1, 07743 Jena, Germany
(Received 26 March 2009; published 17 June 2009)

We present a theory for charge transport in organic crystals. A mobility expression is derived based on the evaluation of the Kubo formula and the Holstein Hamiltonian. It covers the whole temperature range from low T, where it reproduces an expression from the Boltzmann equation for band transport, via elevated T, where it generalizes Holstein's small-polaron theory to finite bandwidths, up to high T, for which a temperature dependence equal to Marcus' electron transfer theory is obtained. The general expression treats coherent band transport and thermally induced hopping on equal footing. By avoiding the approximation of narrow polaron bands the theory allows for the description of large and small polarons.

DOI: 10.1103/PhysRevB.79.235206 PACS number(s): 72.80.Le, 71.38.−k, 72.10.−d, 72.20.Fr

I. INTRODUCTION

Driven by promising technological applications the scientific interest in organic semiconductors is constantly growing. Despite many recent experiments,[1–11] the theoretical interpretation and understanding of several properties remains incomplete. In particular, the charge transport in organic semiconductors often displays very different behavior compared to conventional inorganic semiconductors. Traditional concepts such as band transport which are highly successful in inorganic crystals might not be applicable directly to organic crystals.[12,13] In contrast, new theoretical and computational methods are to be developed that go beyond simple modifications of conventional approaches and account for the distinct properties of charge carriers in organic materials including their interactions with phonons. In particular, polaronic effects on charge transport are stronger in organic semiconductors and cannot be treated as a perturbation. This is due to the complexity of the molecular building blocks and their vibrational degrees of freedom as well as the van der Waals interaction between them. Specific intrinsic material properties are best studied in ordered systems such as organic molecular crystals. The long-range order in such crystals allows for a better understanding because it reduces disorder-related anomalies which might bury important structure-property relationships[14] and chemical trends. The reader is referred to a monograph[15] and a recent review[16] for a more comprehensive overview.

One concept which is extensively used for transport in organic materials is the polaron concept.[17,18] It describes a quasiparticle composed of a charge carrier and a lattice polarization cloud which is bound to this carrier. Thereby, the size is an important measure to characterize polarons in semiconducting organic molecular crystals since Holstein's small-polaron theory or extensions inspired by it are valid for small polarons only.[17,19–22] More precisely, they are restricted to the narrow-band limit where the electronic coupling between neighboring molecules is sufficiently small and the electron-phonon coupling plays a dominant role. In contrast, conventional band theory as known from band transport in inorganic semiconductors[23] is applicable only if the electron-phonon interaction is negligible (no polarons) or at least sufficiently small (large polarons) compared to the electronic coupling. Evidently, these two opposing limits, strong electron-phonon coupling (such as in organic semiconductors) on the one hand and strong electronic coupling (such as in the traditional inorganic semiconductors) on the other, fail to cover the most interesting domain of organic semiconductors with strong electronic coupling. This is an important reason for the incomplete understanding of charge transport in organic crystals. Consequently the transport mechanism is still under debate.[12,15,16] The dilemma gets even more pronounced in the light of the experiments conducted by Warta and Karl.[24] They have demonstrated that in the prototypical naphthalene crystals the transport mechanism may actually depend on the temperature T: band transport at low T and hopping at high T. Moreover, the band transport at low temperatures is not only observed for oligoacenes but it can also be expected for similar crystals of π-conjugate molecules with quite large electronic bandwidths exceeding 500 meV.[11,25,26] In contrast to the bare electronic bandwidth, the polaron bandwidth decreases with growing temperature and becomes much smaller.[27] Therefore, there is a common belief that at high temperatures the polaron becomes localized and is transported by hopping processes.

In the present work, we present a theoretical description of charge transport in organic crystals that unifies the concepts of conventional band theory and polaron hopping in a natural way. We stress that our approach is qualitatively different to the theories based on Holstein's small-polaron model because it describes polarons of arbitrary sizes and, hence, has an extended validity for all temperature. Not only are small polarons included (the high-T limit corresponds to previous narrow-band theories) but also large polarons and even Bloch waves are described in the limit of small and vanishing electron-phonon couplings, respectively. In order to do this, it is *not* necessary to go beyond the Hamiltonian applied previously. We also use the Holstein Hamiltonian but we do not assume narrow bands as in the small-polaron models but, instead, take the full bandwidth into account. As a consequence, a major finding of this work is that we recover the results derived for band transport in inorganic semiconductors (wide bands). In particular, this removes an unphysical low-T singularity which is observed in many previous

treatments[19–22] and which occurs in the conventional treatment of the Holstein model[17] (which initially was not designed for low T) due to the narrow-band approximation. In a previous paper, we have shown that this leads to a significantly better description of the hole mobilities in naphthalene crystals.[28] The goal of the present paper is to give a detailed derivation of this theory including a comparison to previous narrow-band approaches and numerical model studies to illustrate the basic features.

The paper is organized as follows. After this introduction, in Secs. II and III we present the derivation of the theoretical description based upon the Kubo formalism applied to the Holstein Hamiltonian. Hereby, special emphasis is put on the polaron concept and its integration into the calculation of charge-carrier mobilities. An explicit expression for the mobility is given, and the separation into coherent and incoherent transport contributions is discussed in detail. In Sec. IV, we extend our analysis by the discussion of several limiting cases which are included in the theory. Most importantly, we recover the previous narrow-band result (Holstein model) at high T and the conventional wide-band result (Boltzmann equation) for the coherent part of the mobility. These findings are further illustrated by numerical studies in Sec. V. Finally a summary is given in Sec. VI.

II. THEORETICAL DESCRIPTION

A. Basic equations

The experimentally accessible dc mobility of the charge carriers can be extracted from the theoretical quantity current-current correlation function. They are connected by means of the Kubo formula[18]

$$\mu_{\alpha\beta} = \frac{1}{e_0 N_c} \frac{1}{2k_B T} \int_{-\infty}^{\infty} dt \langle j_\alpha(t) j_\beta(0) \rangle_H, \tag{1}$$

which describes the linear current response of a system to an applied electric field. Here, N_c is the number of charge carriers, k_B is Boltzmann's constant, and e_0 is the elementary charge. The mobility tensor $\mu_{\alpha\beta}$ is measurable in time-of-flight experiments or field-effect transistor setups where the charges move on a macroscopic scale beyond their mean-free paths. This corresponds to the validity range of the Kubo formula.

In order to evaluate Eq. (1), the current-current correlation function has to be known for all times t. Hereby, the current operator in Eq. (1) is derived from the polarization operator $\mathbf{P} = e_0 \Sigma_M \mathbf{R}_M a_M^\dagger a_M$ according to

$$\mathbf{j} = \frac{d\mathbf{P}}{dt} = \frac{1}{i\hbar}[\mathbf{P}, H], \tag{2}$$

where the Hamiltonian of the system is the Holstein Hamiltonian[17]

$$H = \sum_{MN} \varepsilon_{MN} a_M^\dagger a_N + \sum_{\mathbf{Q}} \hbar\omega_{\mathbf{Q}} \left(b_{\mathbf{Q}}^\dagger b_{\mathbf{Q}} + \frac{1}{2} \right) + \sum_{MQ} \hbar\omega_{\mathbf{Q}} g_{MM}^{\mathbf{Q}} (b_{\mathbf{Q}}^\dagger + b_{-\mathbf{Q}}) a_M^\dagger a_M. \tag{3}$$

The Hamiltonian, which neglects electron-electron interaction, consists of an electronic part, a phononic part, and a coupling term between electrons and phonons. The particle annihilation (creation) operators $a_M^{(\dagger)}$ are represented in real space (site $\mathbf{R}_M$) and the phonon annihilation (creation) operators $b_{\mathbf{Q}}^{(\dagger)}$ are represented in reciprocal space [mode $\mathbf{Q} \equiv (\mathbf{q}, \lambda)$]. The electron transfer integrals ε_{MN} describe the electronic coupling between orbitals at sites $\mathbf{R}_M$ and $\mathbf{R}_N$, and give rise to finite bandwidth and band dispersion in the periodic systems under study. The on-site energies ε_{MM} can be set to zero according to the freedom of choosing an energy zero. The lattice-vibration part of the Hamiltonian can be assumed to be diagonalized with the frequencies $\omega_{\mathbf{Q}} \equiv \omega_\lambda(\mathbf{q})$ since phonon calculations and diagonalization of the dynamical matrices are common tasks for modern *ab initio* codes. The coupling term in Eq. (3) consists of particle and phonon operators linked by the electron-phonon coupling constants $g_{MM}^{\mathbf{Q}}$. These dimensionless quantities are written in a mixed representation according to the definition of the operators. Similar to the Holstein model, we take the local coupling into account. The inclusion of additional nonlocal electron-phonon interaction $g_{MN}^{\mathbf{Q}}$ has not been considered in the present framework in order to reduce the complexity of the derivation. In addition to the phonon frequencies, the material parameters ε_{MN} and $g_{MM}^{\mathbf{Q}}$ may also be derived from *ab initio* calculations (see, e.g., Refs. 29 and 27).

From Eqs. (2) and (3) it follows that the current operator j_α $(\alpha = x, y, z)$ takes the form

$$j_\alpha = \frac{e_0}{i\hbar} \sum_{MN} (R_{M\alpha} - R_{N\alpha}) \varepsilon_{MN} a_M^\dagger a_N. \tag{4}$$

In order to evaluate the Kubo formula in Eq. (1) the time evolution of $\mathbf{j}$ has to be computed along with thermal averages defined through

$$\langle A \rangle_H = \frac{\mathrm{Tr}(e^{-(H/k_B T)} A)}{\mathrm{Tr}(e^{-(H/k_B T)})}. \tag{5}$$

However, Hamiltonian (3) cannot be diagonalized exactly[18] due to the interaction of particles and phonons and, thus, inhibits a direct evaluation of the thermal average for the current-current correlation function.

B. Polaron transformation

In order to proceed with the evaluation of the correlation function in Eq. (1), we change over to the polaron picture. The canonical transformation

$$\tilde{H} = e^S H e^{S^\dagger}, \tag{6}$$

$$S = \sum_{MQ} g^{\mathbf{Q}}_{MM} a^{\dagger}_{M} a_{M} (b^{\dagger}_{\mathbf{Q}} - b_{-\mathbf{Q}}), \tag{7}$$

which is also known as Lang-Firsov transformation results[30] in the polaron Hamiltonian

$$\tilde{H} = \sum_{MN} a^{\dagger}_{M} \hat{E}_{MN} a_{N} + \sum_{\mathbf{Q}} \hbar \omega_{\mathbf{Q}} \left(b^{\dagger}_{\mathbf{Q}} b_{\mathbf{Q}} + \frac{1}{2} \right), \tag{8}$$

where $\hat{E}_{MN} = e^{C_M} \varepsilon_{MN} e^{-C_N}$ still contains phonon operators through the quantities $C_M = \Sigma_{\mathbf{Q}} g^{\mathbf{Q}}_{MM} (b^{\dagger}_{\mathbf{Q}} - b_{-\mathbf{Q}})$. The same transformation is made for the current operators $\tilde{j}_{\alpha} \equiv e^{S} j_{\alpha} e^{S^{\dagger}}$. With the definition $R_{MN\alpha} \equiv R_{M\alpha} - R_{N\alpha}$ one obtains

$$\tilde{j}_{\alpha} = \frac{e_0}{i\hbar} \sum_{MN} e^{C_M} R_{MN\alpha} \varepsilon_{MN} e^{-C_N} a^{\dagger}_{M} a_{N}. \tag{9}$$

The appearance of additional phonon operators through the quantities C_M reflects the dressing of the bare particles by phonons. Using the transformed operators of Eqs. (8) and (9), the current-current correlation function in Eq. (1) may be exactly rewritten into

$$\langle j_{\alpha}(t) j_{\beta}(0) \rangle_H = \langle e^{iHt/\hbar} j_{\alpha} e^{-iHt/\hbar} j_{\beta} \rangle_H = \langle e^{i\tilde{H}t/\hbar} \tilde{j}_{\alpha} e^{-i\tilde{H}t/\hbar} \tilde{j}_{\beta} \rangle_{\tilde{H}}. \tag{10}$$

C. Polaron concept

At this step it is important to include a brief discussion to clarify some points that, although maybe trivial, appear to get mixed up frequently. First, the polaron transformation as carried out above is made for arbitrary coupling parameters $g^{\mathbf{Q}}_{MM}$. Some authors make use of the term "small-polaron transformation" which, in our eyes, is misleading since the size of the polaron does not depend on the transformation itself. Rather, it depends on the magnitude of the coupling parameters $g^{\mathbf{Q}}_{MM}$. In fact, large polarons (small g) and even fully delocalized Bloch electrons ($g=0$) are also covered by the polaron transformation. In the latter case it trivially holds that $\tilde{H}=H$. Therefore, the Holstein Hamiltonian and the polaron transformation are not *a priori* restricted to small polarons but allow for arbitrary sizes of polarons.

Second, in order to evaluate the Kubo formula [Eq. (1)] analytically, it is necessary to perform a thermal average $\langle \ldots \rangle_H$ as defined in Eq. (5), which requires the diagonalization of H, i.e., a separation between electron and phonon operators. However, the Holstein Hamiltonian H in its general form (3) cannot be diagonalized exactly, and consequently one has to proceed with an approximate diagonalization of H. One idea is to find a hierarchy of energies in the spirit of perturbation theory and neglect small terms. Such a method requires *a priori* assumptions about coupling strengths and, therefore, cannot be of general validity but may give results for limiting cases such as the strong-coupling regime (large g). In this case the electron-phonon interaction is assumed to be larger than the electronic coupling and the latter quantity is treated as a perturbation. Alternatively one considers the weak-coupling regime for small g.

Such considerations, however, are not necessary after the polaron transformation. While the polaron transformation does not exactly diagonalize the Holstein Hamiltonian H (because electron and phonon operators are still coupled through $\hat{E}_{MN}$ which is a complicated function of phonon operators), the two contributions in transformed Hamiltonian (8) can be regarded as a polaronic and a phononic one, which indicates a better route toward decoupling. The advantage of this approach over the separation between electrons and phonons in H is obvious. A separation between polarons (of arbitrary size) and phonons in $\tilde{H}$ is possible without the restriction to either weak-coupling or strong-coupling regime by the replacement of the operator $\hat{E}_{MN}$ by its thermal average over the phononic part of Eq. (3) $\tilde{\varepsilon}_{MN}$. This approach includes the band narrowing and, hence, the variation in the polaron size with temperature. The introduction of the polarons by replacing $\hat{E}_{MN} \rightarrow \tilde{\varepsilon}_{MN}$ has been considered and discussed in previous work[18,31] and results in the approximate polaron Hamiltonian

$$\tilde{H} = \sum_{MN} a^{\dagger}_{M} \tilde{\varepsilon}_{MN} a_{N} + \sum_{\mathbf{Q}} \hbar \omega_{\mathbf{Q}} \left(b^{\dagger}_{\mathbf{Q}} b_{\mathbf{Q}} + \frac{1}{2} \right), \tag{11}$$

which is used henceforth.

The great merit of the polaron transformation is seen from identity (10) which suggests using an approximate diagonalized $\tilde{H}$, which includes the polaron effects, i.e., the coupling constants $g^{\mathbf{Q}}_{MM}$ in all orders, instead of an approximate diagonalized H. In combination with the exact expression (9) for the current operators $\tilde{j}$, the approximate diagonalization of $\tilde{H}$ on top of the polaron transformation is a very general way to account for arbitrary electron-phonon coupling strength as pointed out previously.[31] The new quantities $\tilde{\varepsilon}_{MN}$ are the polaron transfer integrals

$$\tilde{\varepsilon}_{MN} = \varepsilon_{MN} \exp\left[- \sum_{\mathbf{Q}} \left(\frac{1}{2} + N_{\mathbf{Q}} \right) |g^{\mathbf{Q}}_{MM} - g^{\mathbf{Q}}_{NN}|^2 \right], \tag{12}$$

which depend on the electron transfer integrals ε_{MN}, the coupling of the electrons to all phonon modes $\mathbf{Q}$, and the occupation numbers of the phonons $N_{\mathbf{Q}}$, or in other words, on the temperature T. Thereby, $N_{\mathbf{Q}}$ is the occupation number according to the Bose-Einstein statistics $N_{\mathbf{Q}} = (\exp[\frac{\hbar\omega_{\mathbf{Q}}}{k_B T}] - 1)^{-1}$. The polaronic character of $\tilde{\varepsilon}_{MN}$ reduces the bare transfer integrals ε_{MN} according to Eq. (12) and leads to the effect known as band narrowing. The renormalization of the transfer integrals can be rephrased in terms of the temperature-dependent polaron effective mass which is increased by the inverse of the exponential factor in Eq. (12). Equation (12) indicates two contributions to the lowering of the electronic coupling. The first temperature-independent contribution is due to zero-point vibration effects. The second one enters for finite T. Its temperature dependence is due to the phonon occupation $N_{\mathbf{Q}}$ and, hence, increases with rising T. The strength of the narrowing is governed by the magnitude of the electron-phonon coupling constants $g^{\mathbf{Q}}_{MM}$.

We note that for the noninteracting case $g \equiv 0$ the bare quantities ε_{MN} are retained and the Hamiltonian $\tilde{H}$ of Eq.

(11) coincides with the starting Hamiltonian H in Eq. (3). In this limiting case, the effective mass becomes the bare effective mass as obtained from band-structure calculations.

D. Time evolution

We proceed by inserting Eq. (9) into Eq. (10) and obtain

$$\begin{aligned}\langle j_\alpha(t)j_\beta(0)\rangle_H &= \left(\frac{e_0}{i\hbar}\right)^2 \sum_{KLMN} \langle e^{i\tilde{H}t/\hbar} e^{C_K} R_{KL\alpha}\varepsilon_{KL} \\ &\times e^{-C_L} a_K^\dagger a_L e^{-i\tilde{H}t/\hbar} e^{C_M} R_{MN\beta}\varepsilon_{MN} e^{-C_N} a_M^\dagger a_N\rangle_{\tilde{H}} \\ &= \left(\frac{e_0}{i\hbar}\right)^2 \sum_{KLMN} R_{KL\alpha}\varepsilon_{KL} R_{MN\beta}\varepsilon_{MN} \langle e^{C_K(t)} e^{-C_L(t)} \\ &\times e^{C_M} e^{-C_N}\rangle_{\tilde{H}} \langle e^{i\tilde{H}t/\hbar} a_K^\dagger a_L e^{-i\tilde{H}t/\hbar} a_M^\dagger a_N\rangle_{\tilde{H}}. \end{aligned} \quad (13)$$

The second step makes use of the decoupling of polarons and phonons in polaron Hamiltonian (11), and introduces the time evolution of the phonon part via the quantities

$$C_M(t) = \sum_{\mathbf{Q}} g_{MM}^{\mathbf{Q}} (b_{\mathbf{Q}}^\dagger e^{i\omega_{\mathbf{Q}}t} - b_{-\mathbf{Q}} e^{-i\omega_{\mathbf{Q}}t}). \quad (14)$$

The time evolution of the electron operators could be easily computed if the Hamiltonian $\tilde{H}$ would be diagonal also in the operators a and $a^\dagger$, i.e., if the off-diagonal elements $\tilde{\varepsilon}_{MN}$ would vanish. According to Eq. (12), this is the case for high enough temperatures because of the narrowing of the band. In this case one could simply set

$$\tilde{H} \to \tilde{H}' = \sum_M \tilde{\varepsilon}_{MM} a_M^\dagger a_M + \sum_{\mathbf{Q}} \hbar\omega_{\mathbf{Q}} \left(b_{\mathbf{Q}}^\dagger b_{\mathbf{Q}} + \frac{1}{2}\right). \quad (15)$$

In the general case, however, Eq. (15) is an approximation which neglects the bandwidth completely. The Hamiltonian in Eq. (15) is identified with the so-called narrow-band approximation which has been used by many authors in the past.[17,20,22] For strong electron-phonon coupling it gives reasonable results for elevated temperatures but fails completely in the $T \to 0$ limit where it suffers from an unphysical $1/T$ divergence for the mobility. The failure gets even worse if the electron-phonon coupling is not too strong since the narrow-band case is then only fulfilled for extremely high T.

In this paper we follow another route to evaluate Eq. (13) which avoids the narrow-band approximation. The central idea is to take the full bandwidth into account which vastly extends the validity range of the final result. The exact diagonalization of the polaronic part of the Hamiltonian $\tilde{H}$ in Eq. (11) is performed in reciprocal space. We introduce $\mathbf{k}$-space electron creation operators $a_{\mathbf{k}}^\dagger$ and substitute

$$a_M^\dagger = \sqrt{\frac{1}{N_\Omega}} \sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{R}_M} a_{\mathbf{k}}^\dagger, \quad (16)$$

where the sum runs over all wave vectors in the Brillouin zone and N_Ω is the number of unit cells. With the polaron band energies in reciprocal space

$$\tilde{\varepsilon}(\mathbf{k}) = \sum_N \tilde{\varepsilon}_{0N} e^{-i\mathbf{k}\mathbf{R}_N}, \quad (17)$$

polaron Hamiltonian (11) becomes diagonal in this representation

$$\tilde{H} = \sum_{\mathbf{k}} \tilde{\varepsilon}_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{Q}} \hbar\omega_{\mathbf{Q}} \left(b_{\mathbf{Q}}^\dagger b_{\mathbf{Q}} + \frac{1}{2}\right). \quad (18)$$

As a consequence, the time evolution of the electron operators can then easily be computed with the polaron band energy $\tilde{\varepsilon}(\mathbf{k})$ according to

$$e^{(it/\hbar)\tilde{H}} a_{\mathbf{k}}^\dagger e^{-(it/\hbar)\tilde{H}} = a_{\mathbf{k}}^\dagger e^{(it/\hbar)\tilde{\varepsilon}(\mathbf{k})}, \quad (19)$$

and Eq. (13) becomes

$$\begin{aligned}\langle j_\alpha(t)j_\beta(0)\rangle_H = -\left(\frac{e_0}{\hbar}\right)^2 \sum_{KLMN} \frac{1}{N_\Omega^2} \sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} e^{i\mathbf{k}_1\mathbf{R}_K} e^{-i\mathbf{k}_2\mathbf{R}_L} \\ \times e^{i\mathbf{k}_3\mathbf{R}_M} e^{-i\mathbf{k}_4\mathbf{R}_N} e^{(it/\hbar)\tilde{\varepsilon}(\mathbf{k}_1)} e^{-(it/\hbar)\tilde{\varepsilon}(\mathbf{k}_2)} \\ \times \langle a_{\mathbf{k}_1}^\dagger a_{\mathbf{k}_2} a_{\mathbf{k}_3}^\dagger a_{\mathbf{k}_4}\rangle_{\tilde{H}} R_{KL\alpha}\varepsilon_{KL} R_{MN\beta}\varepsilon_{MN} \\ \times \langle e^{C_K(t)} e^{-C_L(t)} e^{C_M} e^{-C_N}\rangle_{\tilde{H}}. \end{aligned} \quad (20)$$

E. Thermal average

Another great advantage of the decoupling of polarons and phonons in polaron Hamiltonian (11) is that electron and phonon operators are fully separated in correlation function (20). For the electron operators, we may calculate the thermal average immediately according to Wick's theorem

$$\langle a_{\mathbf{k}_1}^\dagger a_{\mathbf{k}_2} a_{\mathbf{k}_3}^\dagger a_{\mathbf{k}_4}\rangle_{\tilde{H}} = \delta_{\mathbf{k}_1}^{\mathbf{k}_2}\delta_{\mathbf{k}_3}^{\mathbf{k}_4} n_{\mathbf{k}_1} n_{\mathbf{k}_3} + \delta_{\mathbf{k}_1}^{\mathbf{k}_4}\delta_{\mathbf{k}_3}^{\mathbf{k}_2} n_{\mathbf{k}_1}(1 - n_{\mathbf{k}_2}), \quad (21)$$

where the Fermi-Dirac distribution

$$n_{\mathbf{k}} = \left(\exp\left[\frac{\tilde{\varepsilon}(\mathbf{k}) - \zeta}{k_B T}\right] + 1\right)^{-1}, \quad (22)$$

with the chemical potential $\zeta = \zeta(T, N_c)$ has been introduced.

The two terms on the right-hand side of Eq. (21) are not of equal importance. The n^2 term describes density correlations and does usually not contribute to the transport. In any case it is much smaller than the $n(1-n)$ term if one considers low concentration. Without the n^2 term one obtains

$$\begin{aligned}\langle j_\alpha(t)j_\beta(0)\rangle_H = -\left(\frac{e_0}{\hbar}\right)^2 \sum_{KLMN} \frac{1}{N_\Omega^2} \sum_{\mathbf{k}_1\mathbf{k}_2} e^{i\mathbf{k}_1(\mathbf{R}_K - \mathbf{R}_N)} e^{i\mathbf{k}_2(\mathbf{R}_M - \mathbf{R}_L)} \\ \times e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2)]} R_{KL\alpha}\varepsilon_{KL} R_{MN\beta}\varepsilon_{MN} n_{\mathbf{k}_1} \\ \times (1 - n_{\mathbf{k}_2}) \langle e^{C_K(t)} e^{-C_L(t)} e^{C_M} e^{-C_N}\rangle_{\tilde{H}}. \end{aligned} \quad (23)$$

We still have to evaluate the thermal average over the phonon degrees of freedom. This is much more complicated but can be performed exactly. Here we only describe the major technical steps. One essentially has to calculate the expectation value of products of phonon operators of the type

$$\langle B_{\mathbf{Q}_1}(t) B_{\mathbf{Q}_2}(t) \ldots B_{\mathbf{Q}_i}(t) B_{\mathbf{Q}_1'}(0) B_{\mathbf{Q}_2'}(0) \ldots B_{\mathbf{Q}_j'}(0)\rangle_{\tilde{H}}, \quad (24)$$

where

$$B_{\mathbf{Q}}(t) = b_{\mathbf{Q}}^{\dagger} e^{i\omega_{\mathbf{Q}} t} - b_{-\mathbf{Q}} e^{-i\omega_{-\mathbf{Q}} t}. \tag{25}$$

Since the phonons are interaction-free we make use of the Wick theorem reducing Eq. (24) to all pairwise contractions. Letting r be the number of pairwise contractions with B operators at different times, we can write for the first term with $0 \le r \le \min\{l, j\}$

$$\begin{aligned}\langle B_{\mathbf{Q}_1}(t) B_{\mathbf{Q}_1'}(0)\rangle_{\tilde{H}} \langle B_{\mathbf{Q}_2}(t) B_{\mathbf{Q}_2'}(0)\rangle_{\tilde{H}} \cdots \langle B_{\mathbf{Q}_r}(t) B_{\mathbf{Q}_r'}(0)\rangle_{\tilde{H}} \\ \times \langle B_{\mathbf{Q}_{r+1}}(t) B_{\mathbf{Q}_{r+2}}(t)\rangle_{\tilde{H}} \cdots \langle B_{\mathbf{Q}_{l-1}}(t) B_{\mathbf{Q}_l}(t)\rangle_{\tilde{H}} \\ \times \langle B_{\mathbf{Q}_{r+1}'}(0) B_{\mathbf{Q}_{r+2}'}(0)\rangle_{\tilde{H}} \cdots \langle B_{\mathbf{Q}_{j-1}'}(0) B_{\mathbf{Q}_j'}(0)\rangle_{\tilde{H}}.\end{aligned} \tag{26}$$

Such a complete contraction of Eq. (24) is only nonzero if $l-r$ and $j-r$ are even, and can be evaluated from the thermal averages $\langle B_{\mathbf{Q}}(t) B_{\mathbf{Q}'}(0)\rangle_{\tilde{H}} = -\delta_{\mathbf{Q},-\mathbf{Q}'} \Phi_{\mathbf{Q}}(t)$, where

$$\Phi_{\mathbf{Q}}(t) = N_{\mathbf{Q}} e^{i\omega_{\mathbf{Q}} t} + (1 + N_{\mathbf{Q}}) e^{-i\omega_{\mathbf{Q}} t}. \tag{27}$$

Summing up over all possible contractions in all orders, the phonon part of the thermal average in Eq. (23) can be condensed into the form

$$\begin{aligned}\langle e^{C_K(t)} e^{-C_L(t)} e^{C_M} e^{-C_N}\rangle_{\tilde{H}} = \exp\left[-\frac{1}{2}\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(0) G^{\mathbf{Q}}_{KLKL}\right] \\ \times \exp\left[-\frac{1}{2}\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(0) G^{\mathbf{Q}}_{MNMN}\right] \\ \times \exp\left[-\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(t) G^{\mathbf{Q}}_{KLMN}\right],\end{aligned} \tag{28}$$

where the abbreviation

$$G^{\mathbf{Q}}_{KLMN} = (g^{\mathbf{Q}}_{KK} - g^{\mathbf{Q}}_{LL})(g^{-\mathbf{Q}}_{MM} - g^{-\mathbf{Q}}_{NN}) \tag{29}$$

has been introduced. Note that the first two exponential factors in Eq. (28) are of equal quality. They give rise to the bandwidth narrowing known already from Eq. (12) and reduce the mobility. In contrast, the third exponential factor can amplify the mobility, and explicitly describes phonon absorption and emission events as will be demonstrated later in the paper.

Finally, we collect the above findings, and insert Eqs. (23) and (28) into expression (1) for the mobility tensor. Introducing the short hand notation $\tilde{\varepsilon}_N \equiv \tilde{\varepsilon}_{0N}$ for the polaron transfer integrals, changing summation indices, and summing up over the real-space index K, we obtain

$$\begin{aligned}\mu_{\alpha\beta} = -\frac{1}{e_0 N_c 2 k_B T}\left(\frac{e_0}{\hbar}\right)^2 \sum_{LMN} R_{L\alpha} \tilde{\varepsilon}_L R_{N\beta} \tilde{\varepsilon}_N \frac{1}{N_\Omega} \\ \times \sum_{\mathbf{k}_1 \mathbf{k}_2} e^{-i\mathbf{k}_1(\mathbf{R}_M + \mathbf{R}_N)} e^{i\mathbf{k}_2(\mathbf{R}_M - \mathbf{R}_L)} n_{\mathbf{k}_1}(1 - n_{\mathbf{k}_2}) \\ \times \int_{-\infty}^{\infty} dt e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2)]} \\ \times \exp\left\{-\left[\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(t) G^{\mathbf{Q}}_{0L0N} e^{-i\mathbf{Q}\mathbf{R}_M}\right]\right\},\end{aligned} \tag{30}$$

which is the primary result for the carrier mobility of the present derivation. It includes coherent band transport and incoherent hopping as will be demonstrated in the following.

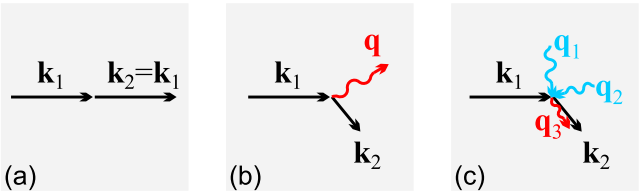

FIG. 1. (Color online) (a) Typical coherent and [(b) and (c)] incoherent processes involved in carrier transport. In (c) a third order process (one emitted and two absorbed phonons) with a resulting phonon wave vector $\mathbf{q} = \mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_3$ is displayed

Importantly, the general form of Eq. (30) allows for the application to arbitrary phonon modes $\omega_{\mathbf{Q}}$ and arbitrary matrix elements $g^{\mathbf{Q}}_{MM}$ of the electron-phonon coupling.

III. CONTRIBUTIONS TO THE MOBILITY

A. Coherent processes

We proceed in order to extract the essential physics and discuss Eq. (30) in terms of contributing scattering events. This is most intuitive if we first split off the zeroth order of electron-phonon interaction in the third line of Eq. (30) according to

$$\begin{aligned}\exp\left[-\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(t) G^{\mathbf{Q}}_{0L0N} e^{-i\mathbf{Q}\mathbf{R}_M}\right] \\ = 1 + \left\{\exp\left[-\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(t) G^{\mathbf{Q}}_{0L0N} e^{-i\mathbf{Q}\mathbf{R}_M}\right] - 1\right\}.\end{aligned} \tag{31}$$

In terms of physics we separate coherent transport (no phonon scattering) from incoherent transport (scattering by phonons). This leads to

$$\mu_{\alpha\beta} = \mu^{(\mathrm{coh})}_{\alpha\beta} + \mu^{(\mathrm{inc})}_{\alpha\beta}. \tag{32}$$

The coherent contribution to the mobility is given as

$$\begin{aligned}\mu^{(\mathrm{coh})}_{\alpha\beta} = -\frac{e_0}{2 N_c k_B T \hbar^2} \sum_{LMN} R_{L\alpha} \tilde{\varepsilon}_L R_{N\beta} \tilde{\varepsilon}_N \frac{1}{N_\Omega} \\ \times \sum_{\mathbf{k}_1 \mathbf{k}_2} e^{-i\mathbf{k}_1(\mathbf{R}_M + \mathbf{R}_N)} e^{i\mathbf{k}_2(\mathbf{R}_M - \mathbf{R}_L)} n_{\mathbf{k}_1}(1 - n_{\mathbf{k}_2}) \\ \times \int_{-\infty}^{\infty} dt e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2)]}.\end{aligned} \tag{33}$$

Mobility expression (33) can be interpreted in terms of contributing scattering events from some initial states $\mathbf{k}_1$ into final states $\mathbf{k}_2$. The sum over $\mathbf{k}_1$ and $\mathbf{k}_2$ in Eq. (33) includes all such events according to the probability $n_{\mathbf{k}_1}$ that an initial state is occupied times the probability $(1 - n_{\mathbf{k}_2})$ of the final state being empty (Pauli blocking factor). From the time integration in Eq. (33) it becomes obvious that the energy of the initial polaron $\tilde{\varepsilon}(\mathbf{k}_1)$ has to match the energy of the final polaron $\tilde{\varepsilon}(\mathbf{k}_2)$ (energy conservation). Moreover, considering the real-space sum over the index M which only occurs in the exponentials gives a Kronecker delta in the wave vectors of initial and scattered polaron $\delta_{\mathbf{k}_1,\mathbf{k}_2}$ (momentum conservation). The momentum conservation [cf. Fig. 1(a)] in the po-

laron wave vectors in $\mu^{(\mathrm{coh})}$ reflects the coherence aspect of band transport, i.e., the moving particle does not loose its phase coherence and its momentum relaxation length is infinite. From the momentum conservation it follows that $\tilde{\varepsilon}(\mathbf{k}_1)=\tilde{\varepsilon}(\mathbf{k}_2)$ is immediately fulfilled and the mobility $\mu^{(\mathrm{coh})}$ becomes infinite, as expected for coherent transport without any scattering mechanism. In real crystals, the coherence can be reduced by other scattering mechanism beyond our model of electron-phonon interaction (impurities, disorder, electron-electron scattering, etc.). Such processes may be accounted for by the introduction of a disorder parameter τ, which reduces the polaron lifetime, as in the literature (see, e.g., Ref. 22). Therefore, we introduce a limiting scattering time from static disorder and describe it by a Gaussian distribution

$$\int dt \to \int dt e^{-(t/\tau)^2}. \tag{34}$$

The broadening limits the coherence time to τ. The corresponding energy can be regarded as a static disorder parameter which mimics different on-site energies, for example.

The resulting mobility can be written as

$$\mu_{\alpha\beta}^{(\mathrm{coh})} = \frac{\sqrt{\pi} e_0 \tau}{2N_c k_B T} \sum_{\mathbf{k}_1} n_{\mathbf{k}_1}(1-n_{\mathbf{k}_1}) \tilde{v}_\alpha(\mathbf{k}_1) \tilde{v}_\beta(\mathbf{k}_1), \tag{35}$$

where the polaron band velocity

$$\tilde{v}_\alpha(\mathbf{k}) = \frac{1}{\hbar} \frac{\partial \tilde{\varepsilon}(\mathbf{k})}{\partial k_\alpha} \tag{36}$$

has been introduced. The fact that the band velocity appears explicitly in Eq. (35) reflects the underlying coherent character of this transport contribution. The tilde reminds us that this quantity is of polaronic nature and therefore temperature dependent because it is affected by the band narrowing described in Eq. (12). Apart from the tildes, Eq. (35) is a well-known expression that may also be derived from the Boltzmann transport equation.[23] However, an essential difference is that expression (35) is now generalized to the transport of polarons instead of bare electrons. In Sec. IV D we discuss the limit of small electron-phonon coupling where the polarons reduce to the bare particles and the expression from the Boltzmann transport equation is recovered.

Finally, we give another form of the result for the coherent part of the mobility. If the squared velocity $\tilde{v}_\alpha^2(\mathbf{k})$ is written as $\tilde{v}_\alpha^2[\tilde{\varepsilon}(\mathbf{k})]$ one can further introduce the polaron density of states (DOS)

$$D(\tilde{\varepsilon}) = \sum_{\mathbf{k}} \delta[\tilde{\varepsilon} - \tilde{\varepsilon}(\mathbf{k})], \tag{37}$$

and write the diagonal elements of the mobility tensor as

$$\mu_\alpha^{(\mathrm{coh})} = \frac{\sqrt{\pi} e_0 \tau}{2N_c k_B T} \int_0^\infty d\tilde{\varepsilon} D(\tilde{\varepsilon}) n_{\tilde{\varepsilon}}(1-n_{\tilde{\varepsilon}}) \tilde{v}_\alpha^2(\tilde{\varepsilon}). \tag{38}$$

Again, Eq. (38) is also similar to a well-known textbook formula for a constant relaxation time[32] with the generalization from electrons to polarons.

B. Incoherent processes

The remaining contribution to the total mobility is due to incoherent-scattering events and reads

$$\begin{aligned}\mu_{\alpha\beta}^{(\mathrm{inc})} = & -\frac{1}{e_0 N_c 2k_B T}\left(\frac{e_0}{\hbar}\right)^2 \sum_{LMN} R_{L\alpha} \tilde{\varepsilon}_L R_{N\beta} \tilde{\varepsilon}_N \frac{1}{N_\Omega} \\ & \times \sum_{\mathbf{k}_1\mathbf{k}_2} e^{-i\mathbf{k}_1(\mathbf{R}_M+\mathbf{R}_N)} e^{i\mathbf{k}_2(\mathbf{R}_M-\mathbf{R}_L)} n_{\mathbf{k}_1} \\ & \times (1-n_{\mathbf{k}_2}) \int_{-\infty}^{\infty} dt e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1)-\tilde{\varepsilon}(\mathbf{k}_2)]} \\ & \times \left(\exp\left\{-\left[\sum_{\mathbf{Q}} \Phi_{\mathbf{Q}}(t) G_{0L0N}^{\mathbf{Q}} e^{-i\mathbf{Q}\mathbf{R}_M}\right]\right\} - 1\right). \end{aligned} \tag{39}$$

Similarly to the preceding paragraph we briefly discuss the scattering events that contribute to the incoherent part of the mobility. First, in Eq. (39) the real-space sum over M leads to the momentum conservation $\delta_{\mathbf{k}_1+\mathbf{q},\mathbf{k}_2}$ with phonon wave vectors. Hereby, $\mathbf{q}$ represents the sum over all phonon vectors which contribute in one scattering process of a certain order of the coupling constants. According to the exponential sum in Eq. (39) this may involve any number of phonons because all orders of electron-phonon interaction contribute. In Fig. 1 incoherent processes of first and third orders are depicted in (b) and (c), respectively. Only in the first order does $\mathbf{q}$ represent the wave vector of a single phonon which scatters the polaron (see also Sec. IV D). Also, the energy conservation is somewhat more complicated compared to the coherent mobility. The energy difference of initial and final polarons is no longer zero as for the coherent part but must account for the phonon energies of created and annihilated vibrations. This is automatically guaranteed via the time integration in Eq. (39). Finally, the weight of such a scattering event is influenced by the occupation numbers $n_{\mathbf{k}_1}(1-n_{\mathbf{k}_2})$ as observed similarly for the coherent contribution. Such incoherent-scattering processes are to be regarded as stochastic processes since the phonon fluctuations are statistically independent. In this way the phase coherence of the particles is destroyed in each scattering event.

If there is no electron-phonon coupling at all, the incoherent contribution completely vanishes. In contrast thereto, a nonzero electron-phonon interaction influences the incoherent contribution through two counteracting effects. The increasing polaron mass with increasing g tends to reduce the mobility. On the other hand, the stronger the electron-phonon coupling the stronger can temperature promote charge transport.

Expression (39) can be simplified if one assumes that the scatterers are dispersionless optical phonons ($\omega_{\mathbf{Q}} \to \omega_\lambda, N_{\mathbf{Q}} \to N_\lambda$), which implies that $\Phi_{\mathbf{Q}}(t) \to \Phi_\lambda(t)$, where

$$\Phi_\lambda(t) = N_\lambda e^{i\omega_\lambda t} + (1+N_\lambda) e^{-i\omega_\lambda t}, \tag{40}$$

and if one, accordingly, sets

$$g_{MM}^{\mathbf{Q}} = \frac{g_\lambda}{\sqrt{N_\Omega}} e^{i\mathbf{q}\mathbf{R}_M}. \tag{41}$$

As a result one obtains

$$\mu_{\alpha\beta}^{(\mathrm{inc})} = -\frac{1}{e_0 N_c 2k_B T}\left(\frac{e_0}{\hbar}\right)^2 \sum_{LMN} R_{L\alpha}\tilde{\varepsilon}_L R_{N\beta}\tilde{\varepsilon}_N \frac{1}{N_\Omega} \times \sum_{\mathbf{k}_1\mathbf{k}_2} e^{-i\mathbf{k}_1(\mathbf{R}_M+\mathbf{R}_N)} e^{i\mathbf{k}_2(\mathbf{R}_M-\mathbf{R}_L)} n_{\mathbf{k}_1} \times (1-n_{\mathbf{k}_2}) \int_{-\infty}^{\infty} dt e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1)-\tilde{\varepsilon}(\mathbf{k}_2)]} \times \left\{ \exp\left[-(\delta_0^M - \delta_L^M - \delta_{-N}^M + \delta_{L-N}^M) \sum_\lambda \Phi_\lambda(t) g_\lambda^2 \right] - 1 \right\}. \tag{42}$$

In order to further simplify the expression, one may concentrate on the most important terms in expression (42). These terms are identified from the minimization of $\delta_0^M - \delta_L^M - \delta_{-N}^M + \delta_{L-N}^M$ in the last exponential. The leading terms are obtained for $L=M$ and $-N=M$. It is confirmed numerically that contributions from all other terms are minor. Taking only the leading terms into account one finds

$$\mu_{\alpha\beta}^{(\mathrm{inc})} = \frac{e_0}{N_c \hbar^2 2k_B T} \sum_L R_{L\alpha} R_{L\beta} \tilde{\varepsilon}_L^2 \frac{1}{N_\Omega} \times \sum_{\mathbf{k}_1\mathbf{k}_2} n_{\mathbf{k}_1}(1-n_{\mathbf{k}_2}) \int_{-\infty}^{\infty} dt e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1)-\tilde{\varepsilon}(\mathbf{k}_2)]} \times \left\{ \exp\left[2\sum_\lambda \Phi_\lambda(t) g_\lambda^2 \right] - 1 \right\} e^{-(t/\tau)^2}. \tag{43}$$

For consistency we have again replaced the time integral according to Eq. (34) introducing the same collision time τ from defect scattering as a disorder parameter. In contrast to the coherent contribution, where the introduction of band velocity (36) was possible, reflecting the delocalization aspect of coherent transport, here the quantity $\frac{1}{\hbar^2}\Sigma_L R_{L\alpha} R_{L\beta} \tilde{\varepsilon}_L^2$ reflects the localization aspect of incoherent transport. This quantity can be looked upon as a measure for the polaron hopping in real space since it is dominated by the nearest-neighbor transfer rates.

IV. LIMITING CASES

A. Narrow-band approximation

In the preceding chapter we have evaluated the mobility from the Kubo formula [Eq. (1)], and obtained an analytical expression [Eq. (30)] which includes coherent and incoherent contributions for arbitrary values of the electron-phonon coupling strength and arbitrary bandwidth. The mobility in Eq. (30) is a generalization of the results of Holstein's small-polaron model because it can also describe the motion of large polarons (small g) and covers even the case of Bloch waves ($g=0$). This was possible because we have avoided the narrow-band approximation, and, instead, incorporated the full bandwidth by means of a mixed real-space and reciprocal-space representation for the electronic and vibronic degrees of freedom. Both contributing transport mechanisms, coherent and incoherent, coexist and have been split for separate discussions. The simplified expressions are given by Eqs. (35) and (43), respectively.

The narrow-band approximation follows from the most general result [Eq. (30)] by setting $\tilde{\varepsilon}(\mathbf{k}_1)=\tilde{\varepsilon}(\mathbf{k}_2)$ for all wave vectors and assuming, accordingly, a constant distribution function $n_{\mathbf{k}} \to c \equiv N_c/N_\Omega$ as well. Subsequent summations over the wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ transform the phase factors to Kronecker deltas in real space. With Eqs. (34), (40), and (41) one finally obtains for Einstein phonons

$$\mu_{\alpha\beta}^{(\mathrm{NBA})} = \frac{e_0(1-c)}{2\hbar^2 k_B T} \sum_L R_{L\alpha} R_{L\beta} \tilde{\varepsilon}_L^2 \int_{-\infty}^{\infty} dt \times \exp\left[2\sum_\lambda \Phi_\lambda(t) g_\lambda^2 \right] e^{-(t/\tau)^2}. \tag{44}$$

This result is the narrow-band result as derived previously for the case of local electron-phonon coupling.[18,22]

The mobility in the narrow-band approximation [Eq. (44)] may also be split into coherent and incoherent parts using the separation $\exp[2\Sigma_\lambda \Phi_\lambda(t) g_\lambda^2] = 1 + (\exp[2\Sigma_\lambda \Phi_\lambda(t) g_\lambda^2] - 1)$, in the spirit of Eq. (31). This is similar to Holstein's original idea of a mobility summation $\mu = \mu^{(1)} + \mu^{(2)}$ reflecting coherent tunneling and hopping motions, respectively.[17] Such a separation also occurs in later theoretical work.[16,27,33] However, in view of the assumed localized nature of the polarons it had not been discussed that $\mu^{(1)}$ should be related to the band transport as obtained from the Boltzmann equation. This is not surprising since the resulting formulas and numerical results for $\mu^{(1)}$ differ strongly from the coherent mobility $\mu^{(\mathrm{coh})}$, Eq. (35), of the present theory that goes beyond the narrow-band limit (see also Sec. V). Consequently, the characteristics of the contribution $\mu^{(1)}$ from such localized quasiparticles in the narrow-band theories appear quite strange. One common flaw is the presence of a low-temperature singularity T^{-1} which does not occur in band transport (see Sec. IV B). This demonstrates that, due to the approximate treatment of the carrier occupation, the narrow-band theories are not valid at lower T, as also previously discussed by Kenkre.[34]

In general, there are two main differences between the present theory and a narrow-band theory. First, in the present approach, there are additional scattering channels from inelastic scattering $[\tilde{\varepsilon}(\mathbf{k}_1) \neq \tilde{\varepsilon}(\mathbf{k}_2)]$, whereas a narrow-band theory takes only elastic scattering into account where the energy difference for initial and final polaron states is always zero $[\tilde{\varepsilon}(\mathbf{k}_1)=\tilde{\varepsilon}(\mathbf{k}_2)]$. Second, the channels which were already included in a narrow-band theory are now much better described since the actual state energy $\tilde{\varepsilon}$ has become a relevant quantity and is incorporated properly. The narrow-band approximation partially neglects the energy dispersion which immediately leads to the incorrect matching of initial and final-state energies.

It is clear that for high enough temperatures the above replacements, which have been introduced to arrive at the narrow-band result, become exact and the full theory coincides with the narrow-band theory. In contrast, for low and medium temperatures we observe strong improvements over the narrow-band approximation. This will be further discussed below and accompanied by numerical studies in Sec. V.

B. Low temperatures

In order to calculate the $T\to 0$ limit of the coherent contribution [Eq. (35)] we make use of the relation $n_{\mathbf{k}}(1-n_{\mathbf{k}})=-k_BT\frac{\partial n_{\tilde{\varepsilon}(\mathbf{k})}}{\partial\tilde{\varepsilon}(\mathbf{k})}$. From the Fermi-Dirac distribution at zero temperature (step function) and its energy derivative, it follows that only the polarons at the chemical potential ζ contribute. The number of contributing polarons from the small energy interval around the chemical potential (thermal layer) decreases in the low-T limit according to the layer width k_BT. As a consequence, the k_BT term exactly cancels the prefactor $\frac{1}{k_BT}$ in Eq. (35) resulting in a finite carrier mobility for $T=0$. Importantly, this is a major improvement over the narrow-band approximation which involves a replacement $n_{\mathbf{k}}(1-n_{\mathbf{k}})\to c(1-c)$ and results in a $\frac{1}{k_BT}$ divergence. Therefore, the correct inclusion of the Fermi-Dirac statistics in the present theory is essential for the removal of this unphysical singularity and constitutes a major advantage of the present theory over the previous narrow-band approach.

In the special case of an isotropic system with a parabolic band structure, we obtain the low-T limit as

$$\mu^{(\mathrm{coh})}=\frac{\sqrt{\pi}e_0\tau}{2m^*_{\mathrm{pol}}}, \tag{45}$$

which resembles the Drude expression for the mobility generalized to polarons as charge carriers. Thereby, the polaron effective mass is given by

$$m^*_{\mathrm{pol}}=m^*_{\mathrm{el/hole}}\exp\left(\sum_{\lambda}g^2_{\lambda}\right), \tag{46}$$

thus, the dressed particle has an enhanced mass with respect to the bare electron/hole due to coupling to the phonons. The occurring exponential factor is the inverse of the factor which occurs for the bandwidth in the $T\to 0$ limit for Einstein phonons [see, e.g., Eq. (12)].

The low-T regime of the incoherent mobility $\mu^{(\mathrm{inc})}$ from Eq. (43) is obtained for a single scatterer λ of frequency ω_λ in the limit of a small phonon occupation number ($N_\lambda\ll 1$). Assuming ultrapure crystals where static disorder is much smaller than phonon induced dynamic disorder, i.e., $1/\tau\ll\omega_\lambda$, we find the leading term

$$\mu^{(\mathrm{inc})}\propto g^2_{\lambda}\frac{N_\lambda}{T}\approx\frac{g^2_{\lambda}}{T}e^{-(\hbar\omega_\lambda/k_BT)}. \tag{47}$$

Basically, this result describes an activation law with the phonon energy as a relevant energy barrier. The process behind this leading term is a phonon absorption process, which elevates the polaron above the Fermi energy. The frequency of occurrence of such an event is proportional to N_λ, the number of available phonons, and the activation law should rather be regarded as a probability than a real barrier. Finally, the $T\to 0$ limit of Eq. (47) is readily obtained as

$$\mu^{(\mathrm{inc})}\to 0, \tag{48}$$

i.e., the incoherent mobility vanishes.

C. High temperatures

For increasing temperature the effective polaron mass strongly increases which is a direct consequence of the band narrowing. Accordingly, we find for the $T\to\infty$ limit that the coherent mobility behaves like

$$\mu^{(\mathrm{coh})}\to 0. \tag{49}$$

For the remaining incoherent contribution, the distribution function for polarons in Eq. (43) becomes constant, $n_{\mathbf{k}}\to c$, if the temperature is high enough that the bandwidth becomes much smaller than k_BT. If in addition the bandwidth is smaller than all relevant phonon energies, the narrow-band limit applies and we can set $\tilde{\varepsilon}(\mathbf{k}_1)=\tilde{\varepsilon}(\mathbf{k}_2)$ in Eq. (43) and obtain

$$\mu^{(\mathrm{inc})}_{\alpha\beta}\to\frac{e_0(1-c)}{2\hbar^2k_BT}\sum_L R_{L\alpha}R_{L\beta}\tilde{\varepsilon}^2_L\int_{-\infty}^{\infty}dt \times\left\{\exp\left[2\sum_{\lambda}\Phi_\lambda(t)g^2_{\lambda}\right]-1\right\}e^{-(t/\tau)^2}. \tag{50}$$

For sufficiently large temperatures ($2N_\lambda g^2_\lambda>1$), the -1 term in the parenthesis is negligibly small and the right-hand side of Eq. (50) equals the narrow-band result from Eq. (44). In fact, Eq. (50) corresponds to the hopping term $\mu^{(2)}$ in Holstein's original narrow-band theory [see our discussion of Eq. (44) in Sec. IV A].

It has been shown in Ref. 22 that in this limiting case, the high-T mobility for a single scatterer λ of frequency ω_λ behaves like

$$\mu^{(\mathrm{inc})}\propto T^{-3/2}e^{-(E_{\mathrm{pol}}/k_BT)}, \tag{51}$$

where $E_{\mathrm{pol}}=\frac{1}{2}g^2_\lambda\hbar\omega_\lambda$ is the polaron binding energy. Since the coherent contribution $\mu^{(\mathrm{coh})}$ vanishes for high T, it follows that this activation law also holds for the total mobility μ in the present theory. Considering the relation of the mobility to the electron transfer rate κ_{ET} according to $\mu\propto\frac{\kappa_{\mathrm{ET}}}{T}$,[16] Eq. (51) has the exact same temperature dependence as obtained in Marcus' electron transfer theory.[35] In the Marcus theory, the polaron binding energy E_p is regarded a barrier for the charge carrier between two states of localization on different sites (initial and final states for the charge-transfer process). Note that beyond the maximum of $\mu(T)$ in Eq. (51) ($k_BT>\frac{2}{3}E_p$), the mobility is a decreasing function of T while for lower temperatures ($k_BT<\frac{2}{3}E_p$) one finds an activation behavior giving larger μ for larger T.

Comparing the high-T limit in Eq. (51) to the low-T limit of the incoherent hopping contribution in Eq. (47), there are differences in the T dependence of the prefactor as well as in the activation energy in the exponential. While at low T the leading term arises from a single phonon-scattering event ($\propto g^2_\lambda$) and is directly related to the number of available phonons N_λ, the high-T limit is caused by the electron-phonon coupling in all orders [since the coupling constant g_λ appears in the exponent in Eq. (51)] and is not easily related to phonon occupation numbers.

D. Small electron-phonon coupling

As an important limiting case, the present theory covers the bare electron limit if one reduces the electron-phonon coupling $g \to 0$ and hence retains the bare electronic band velocity

$$\tilde{v}_\alpha(\mathbf{k}) \to v_\alpha(\mathbf{k}) = \frac{1}{\hbar}\frac{\partial \varepsilon(\mathbf{k})}{\partial k_\alpha}. \tag{52}$$

Accordingly, Eqs. (35) and (38) reduce to

$$\mu_{\alpha\beta}^{(\mathrm{coh})} = \frac{\sqrt{\pi} e_0 \tau}{2 N_c k_B T} \sum_{\mathbf{k}_1} n_{\mathbf{k}_1}(1 - n_{\mathbf{k}_1}) v_\alpha(\mathbf{k}_1) v_\beta(\mathbf{k}_1), \tag{53}$$

and

$$\mu_{\alpha}^{(\mathrm{coh})} = \frac{\sqrt{\pi} e_0 \tau}{2 N_c k_B T} \int_0^\infty d\varepsilon D(\varepsilon) n_\varepsilon (1 - n_\varepsilon) v_\alpha^2(\varepsilon), \tag{54}$$

respectively, where $D(\varepsilon)$ is the ordinary electronic DOS, and $n_{\mathbf{k}}$ and n_ε are taken at the bare electronic structure.

Incoherent transport only contributes for finite electron-phonon interaction. For $g \neq 0$, the electron-phonon coupling in the term $\{\exp[2\Sigma_\lambda \Phi_\lambda(t) g_\lambda^2] - 1\}$ of Eq. (43) enables scattering and thus opens transport channels involving the creation or removal of phonons. For small g we replace $\{\exp[2\Sigma_\lambda \Phi_\lambda(t) g_\lambda^2] - 1\} \to 2\Sigma_\lambda \Phi_\lambda(t) g_\lambda^2$ in Eq. (43) and the time integration, in the limit $\tau \to \infty$, gives rise to delta functions in energy which, for a single scatterer, read

$$\begin{aligned} &\frac{1}{\hbar}\int_{-\infty}^{\infty} dt \{ N_\lambda e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2) + \hbar\omega_\lambda]} \\ &\quad + (1 + N_\lambda) e^{(it/\hbar)[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2) - \hbar\omega_\lambda]} \} \\ &\quad = N_\lambda \delta[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2) + \hbar\omega_\lambda] \\ &\quad + (1 + N_\lambda) \delta[\tilde{\varepsilon}(\mathbf{k}_1) - \tilde{\varepsilon}(\mathbf{k}_2) - \hbar\omega_\lambda]. \end{aligned} \tag{55}$$

From this equation we directly identify both phonon absorption ($\propto N_\lambda$) and emission ($\propto 1 + N_\lambda$) processes with the respective energies $\pm\hbar\omega_\lambda$ that govern the incoherent mobility in the limit of small electron-phonon interaction.

V. NUMERICAL SIMULATIONS

In order to illustrate the difference in the theory with and without narrow-band approximation, and to demonstrate the unique features of the present approach, we supplement our theoretical analysis by some numerical studies. We use an orthorhombic model crystal with a coupling constant $g_\lambda = 1$, a bandwidth of 680 meV which arises from transfer integrals $\varepsilon_a = 100$ meV, $\varepsilon_b = 50$ meV, and $\varepsilon_c = 20$ meV to the nearest neighbors in the three Cartesian directions. The phonon energy for a single scatterer is chosen $\hbar\omega_\lambda = 10$ meV. The lattice constants are $a = 4$ Å, $b = 5$ Å, and $c = 7$ Å. These are considered typical parameters for organic molecular crystals and we use only a single scatterer to simplify the discussion. All calculations were performed with $\hbar/\tau = 0.1$ meV representing an ultrapure crystal.

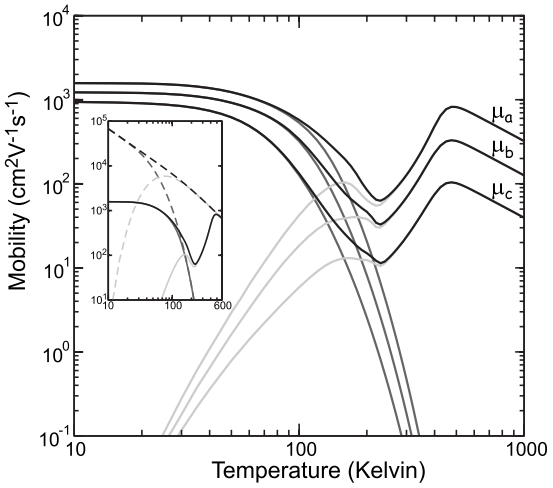

FIG. 2. Hopping mobility $\mu^{(\mathrm{inc})}$ (light gray), band mobility $\mu^{(\mathrm{coh})}$ (dark gray), and total mobility (black) of a model crystal. Different directions are plotted: a, b, and c in order of decreasing mobility. Inset: solid lines represent the present theory (a axis only). Dashed lines show the respective results using the narrow-band approximation.

Figure 2 displays the results for the charge-carrier mobility of the model crystal. The results are plotted for the three Cartesian directions. The observed anisotropy reflects the different lattice vectors and transfer integrals used. For each direction the separation into coherent and incoherent contributions exhibits that the coherent band transport strongly decreases. The hopping contribution shows an activation behavior with rising temperature. Both terms are equal at approximately 160 K and hopping becomes dominant at higher T. Above 250 K the band transport is reduced to a negligible contribution to the total mobility of the model crystal. As a result of the strong decrease in the coherent mobility contribution and the increase in the incoherent contribution, the transition is quite sharp which explains the difficulties of the consistent descriptions within previous theories for one or the other limit.

From the comparison to the narrow-band theory (dashed curves in inset of Fig. 2), we observe that the mobilities in the present approach are lower than the mobilities from the narrow-band theory. In general, both coherent and incoherent mobilities are reduced. At first glance, this result seems to be surprising since the full theory includes both elastic and inelastic elementary processes, whereas the narrow-band theory only accounts for the elastic scattering (see Sec. IV A). This finding in Fig. 2 may be explained by the improved description of the respective scattering events which contribute to the mobility in Eq. (30). In the narrow-band theory, the partial neglect of finite bandwidth effects results in an overestimation of the mobility because the energy conservation ($\tilde{\varepsilon}_1 = \tilde{\varepsilon}_2$) is erroneously taken as fulfilled for all transitions. For finite bandwidth the scattering states are distributed over a broader energy range than in the narrow-band case where all energies are assumed to be equal. In total, this leads to stronger energy mismatch between initial and final scattering states. Applying the correct treatment of energy conservation, such pairs of states with different energies no

longer contribute which leads to a noticeable reduction in the mobility. In the present theory, such a correct description of the energies of initial and final states removes the deficiency from the narrow-band theory.

We also observe a mobility dip at around 250 K. As seen in the inset of Fig. 2, this is traced back to the incoherent contribution. The dip is not present in the narrow-band theory because its physical origin can only be described within the full theory. This is briefly discussed. Starting from the high-T regime with vanishing bandwidth, the bandwidth increases with decreasing temperature. Similar to the above discussion, the reduction in the incoherent mobility with decreasing T is caused by stronger energy mismatch between initial and final states. Below 250 K the incoherent mobility increases again. This is explained by further widening of the bandwidth which enables additional scattering channels (inelastic scattering) once the energy difference $|\tilde{\varepsilon}(\mathbf{k}_1)-\tilde{\varepsilon}(\mathbf{k}_2)|$ exceeds the phonon energy $\hbar\omega_\lambda$. The strongest increase in the mobility is observed at about 200 K which corresponds nicely to a polaron bandwidth of 10 meV. Only below 200 K is this channel for the creation/destruction of phonons with frequency $\hbar\omega_\lambda=10$ meV open and results in higher mobilities.

Another important result is visible in the low-temperature limit of the mobility. While for the narrow-band theory this limit gives an infinite mobility ($\mu\propto 1/T$, see inset of Fig. 2) even if a finite value for τ is chosen, the unphysical low-T singularity is removed in the new approach. This is achieved by taking into account the correct Fermi-Dirac statistics for the particles in the present theory, as discussed in Sec. IV B. As a result, the full theory leads to distinct mobility plateaus for $T\leq 30$ K, which are also observed in experiments on high-quality organic crystals.[24]

Finally, in order to demonstrate the influence of the strength of electron-phonon scattering we varied the coupling parameter g_λ. Figure 3 shows the results for the model crystal with $g_\lambda\in\{0.5,1.0,1.5,2.0\}$. Although the coupling constant is not changed drastically we obtain curves with seemingly different features. While for $g_\lambda=0.5$ the mobility decreases with rising temperature, which is explained by the nearly exclusive contribution of coherent transport up to several hundreds of kelvin, the other curves show both decreasing and increasing total mobilities. This is caused by the band transport-to-hopping transition which shifts to lower temperatures for larger g. This shift can be explained by the reduction in the band transport and promotion of the phonon-assisted hopping with larger g, and is therefore a consequence of the relative size of both contributions. The g dependence for low temperatures is given by Eqs. (45) and (46). For the used values of g we find a modification of the mobilities over nearly two orders of magnitude. The variation is less strong for higher temperatures which can be understood from the compensating effect of the hopping contribution which counteracts the bandwidth narrowing.

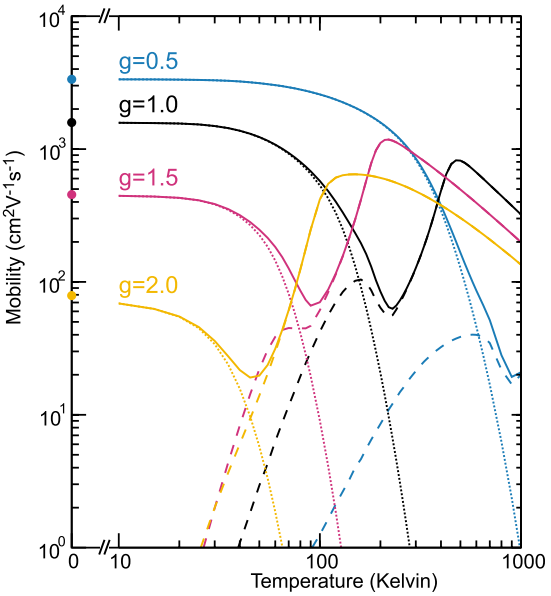

FIG. 3. (Color online) Total mobility (solid lines), and its coherent (dotted lines) and incoherent (dashed lines) contributions along a direction of a model crystal as specified in the text. The electron-phonon coupling is varied as indicated. Dots give the $T=0$ values.

VI. SUMMARY

In summary, we have presented the derivation of a charge transport theory based on a nonperturbative evaluation of the Kubo formula for the carrier mobility which shows several improvements over previous theories. We have been able to overcome the narrow-band approximation commonly used in small-polaron theories and have taken the full bandwidth into account. This extends the validity of the present theory to large polarons and even Bloch waves (which do not couple to the lattice vibrations) as well. The generalization was possible within a mixed real-space and reciprocal-space representation. The discussion of limiting cases has been used to gain additional insight into the theory. In particular, the connection to the results from the Boltzmann equation for low T and the equivalence in the T dependence to the Marcus theory for high T has been demonstrated.

It was found that the mobility can be understood in terms of contributing scattering events, where the inclusion of the full bandwidth allows to take also inelastic-scattering processes into account. The analysis of these scattering events showed that energy and momentum conservation are inherently fulfilled. This analysis led further to the identification of coherent and incoherent contributions to the mobility. While the former shows always a decreasing temperature dependence, the latter is thermally activated for not too high T and a sharp transition between both is observed. Their relative magnitudes, which strongly depends on the strength of the electron-phonon coupling, governs the temperature dependence of the total mobility. The unphysical T^{-1} singularity of narrow-band theories has been removed.

For illustration, we have performed numerical studies on a model crystal. From the comparison of the mobility curves between the present theory and the narrow-band theory the above-mentioned improvements have been demonstrated. In particular the study of the influence of the strength of electron-phonon coupling on the temperature dependence of the mobility showed that a variety of different characteristics can be explained with the theory.

ACKNOWLEDGMENTS

We would like to thank the Deutsche Forschungsgemeinschaft (Project No. HA2900/3-2,3) for financial support. This work was further supported by the European Community within the framework of the Network of Excellence NANOQUANTA (Contract No. NMP4-CT-2004-500198). Grants for computer time from the Leibniz-Rechenzentrum München and the Höchstleistungsrechenzentrum Stuttgart are gratefully acknowledged.

[1] A. J. Heeger, Solid State Commun. **107**, 673 (1998).
[2] H. Sirringhaus, N. Tessler, and R. H. Friend, Science **280**, 1741 (1998).
[3] S. E. Shaheen, C. J. Brabec, N. S. Sariciftci, F. Padinger, T. Fromherz, and J. C. Hummelen, Appl. Phys. Lett. **78**, 841 (2001).
[4] R. W. I. de Boer, T. M. Klapwijk, and A. F. Morpurgo, Appl. Phys. Lett. **83**, 4345 (2003).
[5] V. Y. Butko, X. Chi, D. V. Lang, and A. P. Ramirez, Appl. Phys. Lett. **83**, 4773 (2003).
[6] O. D. Jurchescu, J. Baas, and T. T. M. Palstra, Appl. Phys. Lett. **84**, 3061 (2004).
[7] V. C. Sundar, J. Zaumseil, V. Podzorov, E. Menard, R. L. Willett, T. Someya, M. E. Gershenson, and J. A. Rogers, Science **303**, 1644 (2004).
[8] I. N. Hulea, S. Fratini, H. Xie, C. L. Mulder, N. N. Iossad, G. Rastelli, S. Ciuchi, and A. F. Morpurgo, Nature Mater. **5**, 982 (2006).
[9] A. K. Tripathi and J. Pflaum, Appl. Phys. Lett. **89**, 082103 (2006).
[10] N. Koch, A. Vollmer, I. Salzmann, B. Nickel, H. Weiss, and J. P. Rabe, Phys. Rev. Lett. **96**, 156803 (2006).
[11] G. Koller, S. Berkebile, M. Oehzelt, P. Puschnig, C. Ambrosch-Draxl, F. P. Netzer, and M. G. Ramsey, Science **317**, 351 (2007).
[12] M. E. Gershenson, V. Podzorov, and A. F. Morpurgo, Rev. Mod. Phys. **78**, 973 (2006).
[13] A. Troisi and G. Orlandi, Phys. Rev. Lett. **96**, 086601 (2006).
[14] F. Ortmann, K. Hannewald, and F. Bechstedt, J. Phys. Chem. B **113**, 7367 (2009).
[15] E. A. Silinish and V. Čápek, *Organic Molecular Crystals* (American Institute of Physics, New York, 1994).
[16] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Brédas, Chem. Rev. (Washington, D.C.) **107**, 926 (2007).
[17] T. Holstein, Ann. Phys. **8**, 343 (1959).
[18] G. D. Mahan, *Many-Particle Physics* (Kluwer Academic Publishers, New York, 2000).
[19] R. J. Silbey and R. W. Munn, J. Chem. Phys. **72**, 2763 (1980).
[20] V. M. Kenkre, J. D. Andersen, D. H. Dunlap, and C. B. Duke, Phys. Rev. Lett. **62**, 1165 (1989).
[21] S. Fratini and S. Ciuchi, Phys. Rev. Lett. **91**, 256403 (2003).
[22] K. Hannewald and P. A. Bobbert, Phys. Rev. B **69**, 075212 (2004).
[23] J. M. Ziman, *Electrons and Phonons* (Clarendon Press, Oxford, 2001).
[24] W. Warta and N. Karl, Phys. Rev. B **32**, 1172 (1985).
[25] F. Ortmann, K. Hannewald, and F. Bechstedt, Phys. Rev. B **75**, 195219 (2007).
[26] F. Ortmann, K. Hannewald, and F. Bechstedt, J. Phys. Chem. B **112**, 1540 (2008).
[27] F. Ortmann, K. Hannewald, and F. Bechstedt, Appl. Phys. Lett. **93**, 222105 (2008).
[28] F. Ortmann, K. Hannewald, and F. Bechstedt (unpublished).
[29] K. Hannewald and P. A. Bobbert, Appl. Phys. Lett. **85**, 1535 (2004).
[30] An additional term appears which describes electron-electron interaction mediated by phonons and may be reduced to $-\Sigma_{M\mathbf{Q}} a_M^\dagger a_M |g_{MM}^{\mathbf{Q}}|^2 \hbar\omega_{\mathbf{Q}}$ in accordance to the neglect of explicit electron-electron interaction from the beginning (see also discussion in Ref. 31). This term describes the reduction in the on-site energies by the polaron binding energy and is irrelevant for transport because it shifts only the energy zero.
[31] K. Hannewald, V. M. Stojanović, J. M. T. Schellekens, P. A. Bobbert, G. Kresse, and J. Hafner, Phys. Rev. B **69**, 075211 (2004).
[32] D. K. Ferry, *Semiconductor Transport* (Taylor & Francis, London, 2000).
[33] Y. C. Cheng and R. J. Silbey, J. Chem. Phys. **128**, 114713 (2008).
[34] V. M. Kenkre, Phys. Lett. A **305**, 443 (2002).
[35] R. A. Marcus, Rev. Mod. Phys. **65**, 599 (1993).