Barrier crossing to the small Holstein polaron regime

Peter Hamm[1] and G. P. Tsironis[2,3]

[1]*Physikalisch-Chemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[2]*Department of Physics, University of Crete, P. O. Box 2208, 71003 Heraklion, Crete, Greece*
[3]*Institute of Electronic Structure and Laser, Foundation for Research and Technology–Hellas (FORTH), 71110 Heraklion, Crete, Greece*

We investigate the dimensionality effects of the Holstein polaron from the fully quantum regime, where the crossover between large and small polaron solutions is known to be continuous in all dimensions, into the limit described by the semiclassical discrete nonlinear Schrödinger (DNLS) equation, where the crossover is continuous in one dimension (1D) but discontinuous in higher dimensions. We use exact numerics on one hand and a two variable parametrization of the Toyozawa ansatz on the other in order to probe the crossover region in all parameter regimes. We find that a barrier appears also in 1D separating the two types of solutions, seemingly in contradiction to the common paradigm for the DNLS according to which the crossover is barrier-free. We quantify the polaron behavior in the crossover region as a function of the exciton overlap and find that the barrier remains small in 1D and tunneling through it is not rate-limiting.

The Holstein model has been used widely to describe transport properties of electrons or excitons in diverse systems, ranging from molecular crystals[1] to strongly correlated electron-phonon systems,[2] interface charge localization in alkane layers,[3] organic transistors,[4,5] proteins and proteinlike crystals[6–8] as well as DNA.[9] In the context of the model, the electronic degrees of freedom are coupled to dispersionless phonons leading to the formation of large, i.e., much larger than the lattice spacing, or small, viz. essentially single-site, polarons. The relative magnitudes of three parameters of the theory, viz. electronic overlap integral, phonon energy and exciton-phonon coupling determine the specific polaron properties. Outstanding issues have been the nature of polarons in different regimes, their transitions as well as the precise onset of self-trapping. The discrete nonlinear Schrödinger (DNLS) equation, which is a semiclassical approximation to the Holstein model, gives a continuous transition in one dimension (1D) and a discontinuous transition in higher dimensions,[10,11] while most variational calculations, in particular those based on the Toyozawa wave function, show a discontinuity in all dimensions.[12–18] On the other hand, it has been shown recently with the help of a numerically exact solution of the full-quantum Holstein polaron that the crossover is continuous in 1D (Ref. 19) as well as in higher dimensions.[20,21] The nonexistence of phase transitions in polaron systems can also be proven on very general grounds.[22] In the present Brief Report, we address the nature of the polaron transitions using exact numerics for the fully quantum Holstein model on one hand, as well as a simple, physically motivated two-parameter variational minimization on the other hand. The comparison of the two approaches resolves the issue of the large-to-small polaron transition, produces quantitative information on the nature of the self-trapping transition, and finalizes pending issues regarding deficiencies of standard approximations.

The Holstein Hamiltonian reads

$$H = H_{\mathrm{ex}} + H_{\mathrm{ph}} + H_{\mathrm{ex,ph}},$$

$$H_{\mathrm{ex}} = -J\sum_j (B_j^\dagger B_{j+1} + B_j^\dagger B_{j-1}),$$

$$H_{\mathrm{ph}} = \hbar\omega\sum_j (b_j^\dagger b_j + 1/2),$$

$$H_{\mathrm{ex,ph}} = -\chi\sum_j B_j^\dagger B_j(b_j^\dagger + b_j), \tag{1}$$

where $b_j^\dagger$ (b_j) and $B_j^\dagger$ (B_j) destroy (create) a phonon and an electron (or exciton) at site j, respectively, and J, $\hbar\omega$, χ are the excitonic overlap, phonon energy and exciton-phonon coupling, respectively. In what follows we use two dimensionless parameters, i.e., the exciton coupling $J/\hbar\omega$ and exciton-phonon coupling $\chi/\hbar\omega$ while all energies are given in units of one phonon-quantum $\hbar\omega$. We also restrict our study to the wave number $k=0$ case since this will be the polaron ground state, and because it is most relevant in optical spectroscopy ($k=0$ selection rule).

We first present numerically exact results in all three dimensions using the approach of Trugman and co-workers.[19,20] Figure 1, top row, shows the reduced phonon density at the site of the exciton as a function of the phonon coordinate $q_0 \equiv (b_0^\dagger + b_0)$, with phonon coordinates at sites different from the exciton site traced out. We note that the results are qualitatively similar in all dimensions: The reduced phonon density is centered at $q_0 \approx \chi$ for small exciton couplings J (i.e., a small polaron), similar to the $J=0$ case, for which an exact (analytical) solution exists revealing with $q_0 = \chi$.[6] For sufficiently large exciton couplings J, the phonon displacement shifts to $q_0 \approx 0$, i.e., essentially a free exciton without any phonon displacement. We observe that the transition between the two regimes is smoother in 1D than it is in two dimensions (2D) and three dimensions (3D), however, the essential point is that free exciton and small polaron solutions *coexist* in a certain parameter range, also in 1D. The two states gradually change their relative weights, but hardly their character, as the exciton coupling J is increased. If we interpret the reduced phonon density as one that originates from an effective potential $V(q_0)$, then the coexistence of two solutions hints to the presence of a barrier separating them. This conclusion is in disagreement with the semiclassical DNLS case,[10,11] which does not reveal any barrier in 1D but rather a gradual transition.
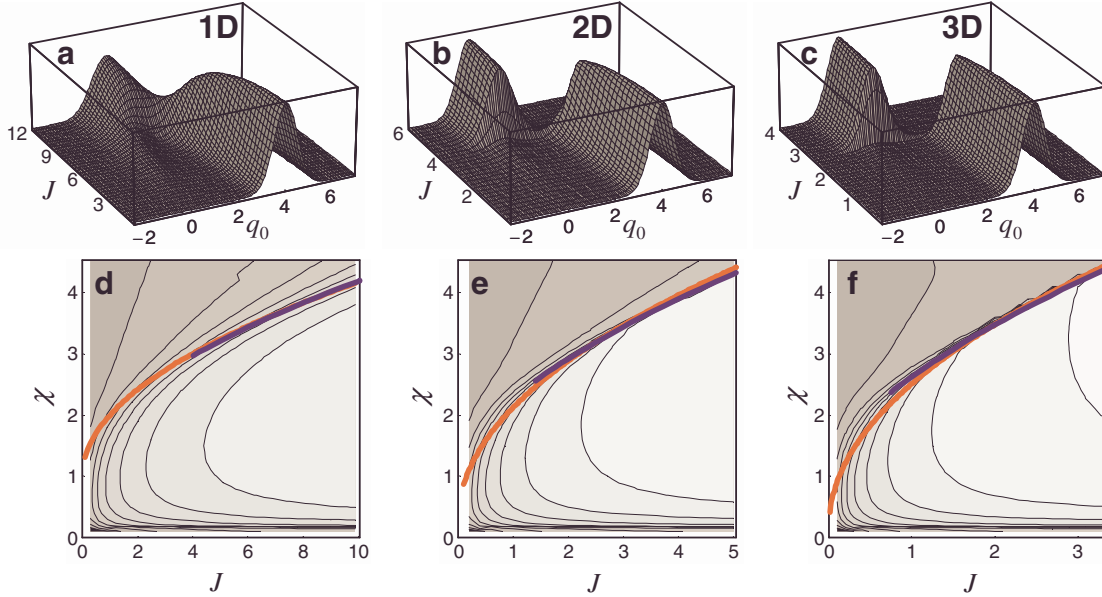
FIG. 1. (Color online) Results from a numerically exact diagonalization of the Holstein Hamiltonian in 1D (left), 2D (middle), and 3D (right). Top row (a)–(c): Phonon density of the polaron ground state as a function of the phonon coordinate q_0. Shown is the reduced phonon density at the site of the exciton. The exciton-phonon coupling was set to $\chi = 4.0$. Bottom row (d)–(f): Phase diagram, plotting the expectation value of the phonon coordinate $\langle q_0 \rangle$ in units of χ for the polaron ground state. The contour lines are equidistant with spacing 0.1. In red (light gray) are the phase separation lines from Eq. (2), whereas in blue (dark gray) are those obtained from equaling the two variational minima of the Toyozawa wave function (see Fig. 2).

The phase diagrams in Fig. 1, bottom row, summarize these results where we plot the expectation value of the mean displacement of the phonon coordinate $\langle q_0 \rangle$ in units of χ (which is the maximum value obtained in the $J=0$ case). For small couplings χ and J, the result is qualitatively very similar in all dimensions with a continuous transition between small and large polaron solutions. Nevertheless, it is quite evident that the transition becomes increasingly more abrupt in 2D and 3D for large couplings (i.e., as we approach the regime of the DNLS), whereas the crossover stays smooth in the 1D case. The phase separation lines can be fitted extremely well to a generic relationship (Fig. 1, bottom row, red lines):

$$\chi_c = 1/\alpha + \sqrt{\alpha J}, \tag{2}$$

with $\alpha = 1$ in 1D, $\alpha = 3.34$ in 2D, and $\alpha = 5.41$ in 3D, respectively. The 1D value stems from Lindenberg *et al.*[23] empirical relationship, whereas those for 2D and 3D coincide with the critical couplings in the DNLS limit (when $\sqrt{\alpha J} \gg 1/\alpha$).[11] Hence, Eq. (2) does give the correct semiclassical limits expected from the DNLS analysis.

Although exact, the results of Fig. 1 originate from a diagonalization of a huge matrix and provide relatively little physical insight. We therefore employ an approximate, yet more intuitive trial function, the Toyozawa wave function,[12–14,17] and use the numerically exact solution as an accuracy check. The Toyozawa trial function is written as a Bloch state:

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_l e^{ikl} |\psi_l\rangle, \tag{3}$$

with

$$|\psi\rangle \equiv \sum_i a_i B_i^\dagger |0\rangle_{\text{ex}} \prod_j |q_{j-i}\rangle_j, \tag{4}$$

where all $|\psi_l\rangle \equiv |\psi\rangle$ are identical, $|0\rangle_{\text{ex}}$ is an exciton vacuum state, and $|q\rangle_j$ is a coherent phonon state:

$$|q\rangle_j \equiv e^{q(b_j^\dagger - b_j)} |0\rangle_{\text{ph}}. \tag{5}$$

The Toyozawa wave function can be considered a Bloch extension of symmetry breaking soliton solutions,[14] where both excitons and phonons are dressing one common trapping site l. In 1D, the norm of the wave function $|\psi\rangle$ is[13]

$$\langle \psi | \psi \rangle = \sum_i \left[\sum_j a_j a_{j-i} \prod_k e^{-(q_k - q_{k-i})^2/2} \right] \tag{6}$$

and the expectation value of the Hamiltonian,

$$\langle \psi | H | \psi \rangle = -2J \sum_i \left[\sum_j a_j a_{j-i} \prod_k e^{-(q_k - q_{k-i+1})^2/2} \right]$$
$$+ \sum_i \left[\sum_j a_j a_{j-i} \sum_l q_l q_{l-i} \prod_k e^{-(q_k - q_{k-i})^2/2} \right]$$
$$- \chi \sum_i \left[\sum_j a_j a_{j-i} (q_j + q_{j-i}) \prod_k e^{-(q_k - q_{k-i})^2/2} \right]. \tag{7}$$

In more than 1D (dimension D), the indices i and j are replaced by $\{i_x, i_y(, i_z)\}$ and $\{j_x, j_y(, j_z)\}$, respectively, in either sum/product. Furthermore, the exciton term is repeated D times with the one-site shift (i.e., the q_{j-i+1}-term) in either direction (for isotropic exciton coupling, as considered here, this can be reduced to one exciton term with prefactor $2DJ$).

Zhao *et al.*[14] have minimized the energy of the Toyozawa wave function by effectively varying all parameters $\{a_i\}$ and
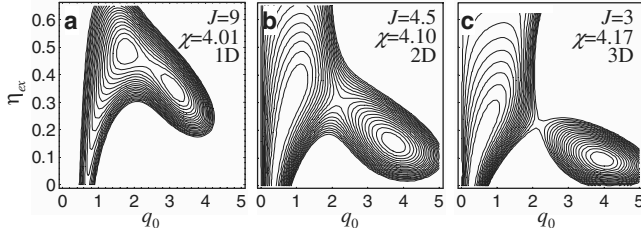
FIG. 2. Variational energy of the Toyozawa wave function as a function of q_0 and η_{ex} in 1D (left), 2D (middle), and 3D (right). Exciton coupling is $J=9/D$ and the exciton-phonon coupling $\chi = \chi_c^{var}$ has been calculated so that both variational minima are equally deep. Contour line spacings are 0.05 in (a) and 0.1 in (b) and (c).

$\{q_i\}$ (the calculation was done in momentum space), still leaving us with a multiparameter solution. In contrast, we make an exponential ansatz, in analogy to Refs. 11, 13, and 17:

$$a_i \equiv \eta_{ex}^{|i|},$$

$$q_i \equiv q_0 \eta_{ph}^{|i|}, \qquad (8)$$

that contains only three parameters. Minimizing Eq. (7) reveals $\Sigma_i q_i = \chi$ as one condition[13,14] which allows us to eliminate one of these parameters leading to $\eta_{ph} = (\sqrt[D]{\chi} - \sqrt[D]{q_0})/(\sqrt[D]{\chi} + \sqrt[D]{q_0})$. Hence, we can express the variational energy as a function of effectively two parameters, q_0 and η_{ex}.

Figure 2 shows the variational energy of the trial function for $J=9/D$. We observe the emergence of two variational minima in all dimensions with a barrier separating them. The barrier is very shallow in 1D (≈ 0.1) while it is much more pronounced in 2D or 3D. This is the barrier which is responsible for the double-peak structure of the wave function in the numerically exact solution, which also exists in 1D [Fig. 1(a)]. Venzl and Fischer[13] have shown a very similar variational energy surface in 1D (as a function of η_{ex} and η_{ph}, eliminating q_0); however, they did not address the double minimum due to a choice of too small value of the exciton coupling, viz. $J=2$.

For the chosen values of the exciton-phonon coupling the two minima are equally deep. Using the degeneracy of the variational minima as a criterion for the critical exciton-phonon coupling χ_c^{var} that induces the large-to-small polaron transition, we find that the numerically obtained line $\chi_c^{var}(J)$ follows very closely the relationship of Eq. (2) [compare blue and red lines in Figs. 1(d) and 1(f)]. This near identity of χ_c^{var} and Eq. (2) starts to deviate in 1D for very large exciton overlaps $J \gtrsim 100$, indicating that Eq. (2) is a lower order expansion of a more complex functional relationship.

Figure 3 shows the energy of the lowest energy eigenstate, as deduced from the numerically exact solution (black line), and compares it with the two variational minima (where they exist, red lines). Outside the crossing region, the variational result agrees reasonably well with the exact energy—hence the two-parameter trial function Eq. (8) is quite good in this regime. However, both deviate significantly at the crossing
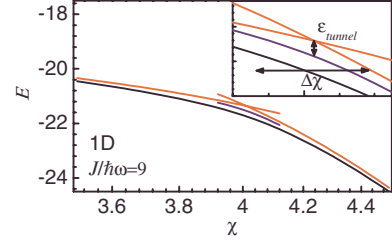

FIG. 3. (Color online) Energies of the lowest energy state for $J=9$ and the exciton-phonon coupling χ varied through the crossover region in 1D. Black line: numerically exact result. Red (light gray) lines: The two minima of the variational energy. Blue (dark gray) line: Energy corrected by tunneling between the two components. The inset focuses into the crossing region and defines the terms introduced in Eq. (10).

point, indicating that the Toyozawa wave function *per se* does not capture the physics of the crossover correctly.[12,14,15] In particular, the variational minimum, as well as all other properties of the trial wave function, would *not* be analytical at the cross-over point. This is simply due to the fact that the two variational minima are separated by a barrier [Fig. 2(a)], and the solution of an overall minimization switches abruptly when one of these local minima decreases below the other.

Barišić has extended the variational space by considering a linear combination of two (or more) Toyozawa wave functions:[17]

$$|\psi\rangle \equiv c_l|\psi^{(l)}\rangle + c_s|\psi^{(s)}\rangle, \qquad (9)$$

showing that this removes the discontinuity in 1D. Figure 2(a) makes very clear why that is so: The two Toyozawa wave functions will sit in either of two minima of the variational energy, and tunneling coupling between them will (a) lower the energy and thereby correct for approximately half of the mismatch to the numerically exact result, and (b) render the crossover continuous (see Fig. 3, blue line).[24]

In order to understand the nature of the barrier separating the small and large polaron solutions we address the question of how the crossover domain behaves in the limit $J \to \infty$. Mixing of the large and the small polaron states will occur whenever the tunnel coupling is on the order of the energy separation between them, $\varepsilon_{tunnel} \gtrsim |E^{(l)} - E^{(s)}|$. Hence, the quotient of the tunnel coupling divided by the difference of the derivatives of the two variational energies with respect to χ defines a measure of the extension $\Delta\chi$ of the crossover region (Fig. 3, inset, shows pictorially ε_{tunnel} and $\Delta\chi$):

$$\Delta\chi \equiv 2\frac{\varepsilon_{tunnel}}{d(E^{(l)} - E^{(s)})/d\chi}. \qquad (10)$$

Figure 4(a), red line, shows that the tunnel coupling ε_{tunnel} stays essentially constant in 1D for all J's, however, the denominator in Eq. (10) decreases with J so that $\Delta\chi$ increases roughly linearly for large enough J [Fig. 4(b), red line]. Consequently, the crossover remains continuous even for $J \to \infty$, leading to the correct 1D result in the adiabatic limit. In contrast, in higher dimensions, both the tunnel coupling ε_{tunnel} and $\Delta\chi$ decay exponentially with J (Fig. 4, blue and green lines). Hence, the superposition of two Toyozawa
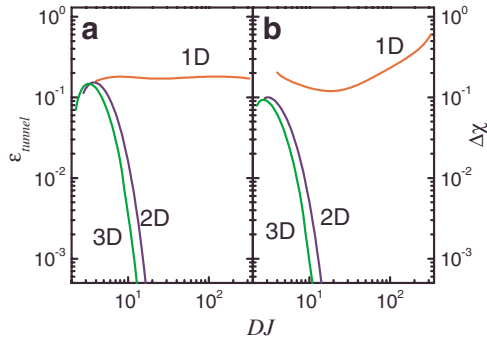
FIG. 4. (Color online) (a) Tunnel coupling $\varepsilon_{\text{tunnel}}$ and (b) extension of the cross-over region $\Delta\chi$ as a function of exciton coupling J (with $\chi=\chi_c^{\text{var}}$) in 1D (red), 2D (blue), and 3D (green).

wave functions, although predicting in principle a continuous crossover, leads (in 2D, 3D) to an increasingly more abrupt transition for $J\rightarrow\infty$, just like the numerically exact result does [Figs. 1(e) and 1(f)].

In conclusion, we studied the dependence of quantum polaron regimes on lattice dimensionality employing exact numerics and a physically motivated variational ansatz based on the Toyozawa wave function. The exact numerical solution shows that the large-to-small polaron crossover is continuous in all dimensions although much more abrupt in two and three dimensions, especially as the parameters approach the adiabatic regime [Figs. 1(e) and 1(f)]. In 1D, the transition is smooth and remains continuous also for large J's [Fig. 1(d)].

More physical insight into the polaron features is obtained from an approximate trial wave function, viz. the Toyozawa wave function of Eq. (4), that leads to variational results whose regime of validity has been analyzed extensively.[14,15] When the latter is augmented with an exponential ansatz of Eq. (8) containing effectively only two variational parameters, it leads to a very intuitive and clear picture. Despite the simplicity of this approach, it provides results that coincide with the exact ones outside the crossover region, while, nevertheless, predicting nonanalytic properties at the transition point in all dimensions. The appearance of this discontinuity was pointed out previously[17,20] but it has been found in this work to be due to the presence of two, almost degenerate solutions coexisting in a certain parameter regime while separated through a barrier. Consequently, a superposition of two Toyozawa wave functions, as in Eq. (9), reveals qualitatively correct results for the cross-over region in all dimensions and in all parameter regimes.[17] The resulting intuitive picture may be used to assess quantitatively the properties of the cross-over region as a function of the exciton coupling. One important outcome of this analysis is that a barrier separating large and small polaron solutions exists also in 1D, in sharp contrast to the predictions of semiclassical theories. We nevertheless find that tunneling through the barrier remains efficient in all parameter regimes [Fig. 4(a), red line], since the barrier stays on the order of $\hbar\omega$ and hence does not hamper motion in the DNLS limit where phonons are treated classically. The Holstein model is a fundamental "minimal" model for strongly interacting systems and thus our results may also have ramifications for more complex models such as for instance the Holstein-Hubbard model.

[1] T. Holstein, Ann. Phys. (N.Y.) **8**, 325 (1959).

[2] W. Z. Wang, A. R. Bishop, J. T. Gammel, and R. N. Silver, Phys. Rev. Lett. **80**, 3284 (1998).

[3] N. H. Ge, C. M. Wong, R. L. Lingle, Jr., J. D. McNeill, K. J. Gaffney, and C. B. Harris, Science **279**, 202 (1998).

[4] L. Torsi, A. Dodabalapur, L. J. Rothberg, A. W. P. Fung, and H. E. Katz, Science **272**, 1462 (1996).

[5] W. A. Schoonveld, J. Wildeman, D. Fichou, P. A. Bobbert, B. J. van Wees, and T. M. Klapwijk, Nature (London) **404**, 977 (2000).

[6] A. C. Scott, Phys. Rep. **217**, 1 (1992).

[7] J. Edler, P. Hamm, and A. C. Scott, Phys. Rev. Lett. **88**, 067403 (2002).

[8] P. Hamm and G. P. Tsironis, Eur. Phys. J. Spec. Top. **147**, 303 (2007).

[9] S. S. Alexandre, E. Artacho, J. M. Soler, and H. Chacham, Phys. Rev. Lett. **91**, 108105 (2003).

[10] V. V. Kabanov and O. Y. Mashtakov, Phys. Rev. B **47**, 6060 (1993).

[11] G. Kalosakas, S. Aubry, and G. P. Tsironis, Phys. Rev. B **58**, 3094 (1998).

[12] Y. Toyozawa, Prog. Theor. Phys. **26**, 29 (1961).

[13] G. Venzl and S. F. Fischer, Phys. Rev. B **32**, 6437 (1985).

[14] Y. Zhao, D. W. Brown, and K. Lindenberg, J. Chem. Phys. **107**, 3159 (1997).

[15] A. H. Romero, D. W. Brown, and K. Lindenberg, J. Chem. Phys. **109**, 6540 (1998).

[16] V. Cataudella, G. De Filippis, and G. Iadonisi, Phys. Rev. B **60**, 15163 (1999).

[17] O. S. Barišić, Phys. Rev. B **65**, 144301 (2002).

[18] O. S. Barišić, Europhys. Lett. **77**, 57004 (2007).

[19] J. Bonča, S. A. Trugman, and I. Batistić, Phys. Rev. B **60**, 1633 (1999).

[20] L.-C. Ku, S. A. Trugman, and J. Bonča, Phys. Rev. B **65**, 174306 (2002).

[21] L.-C. Ku and S. A. Trugman, Phys. Rev. B **75**, 014307 (2007).

[22] B. Gerlach and H. Löwen, Phys. Rev. B **35**, 4291 (1987).

[23] A. H. Romero, D. W. Brown, and K. Lindenberg, Phys. Rev. B **59**, 13728 (1999).

[24] The expansion coefficients c_l and c_s in Eq. (9) are determined by minimization through direct matrix diagonalization. The overlap and mixing matrix elements are calculated in analogy to Eqs. (6) and (7), respectively, with one coefficient a_j or q_k in either term stemming from $|\psi^{(l)}\rangle$, and the other, a_{j-i} or q_{k-i}, from $|\psi^{(s)}\rangle$. Furthermore, in the 1D case, we have allowed the positions of the two Toyozawa wave functions to deviate slightly from the exact minima during the minimization, thereby increasing the mixing matrix element on the cost of increasing the diagonal matrix elements (energies).