The Holstein polaron:

results from numerical and analytical approaches

Collaborators: Glen Goodvin, Lucian Covaci, Dominic Marchand + George Sawatzky and Bayo Lau

Numerical data: A. Macridin, A. Mishchenko, N. Nagaosa, V. Cataudella, P. Kornilovitch

More details: → "Numerical solution of the Holstein polaron problem" by H. Fehske and S. A. Trugman → 70 pages book review, arXiv:cond-mat/0611020

→ Our analytical work on the Momentum Average = MA approximation, see my webpage, www.phas.ubc.ca/~berciu

Funding: Sloan Foundation, ClfAR Nanoelectronics, CFI, NSERC

Polaron = electron + lattice distortion (phonon cloud) surrounding it

and some physics.

→ very old problem: Landau, 1933 > very many models to study, e.g. □ single polaron (one extra charge carrier in an insulator) vs. bi-polarons = bound state of two polarons, vs. many-polarons systems (→ metals, superconductors) \square large polarons (continuous approx) vs. small polarons (lattice model \rightarrow different lattices with d=1,2,3, different couplings, etc.) □ coupling to acoustic or to optical phonon modes, or to both? and then: spin polarons, Jahn-Teller/orbital polarons, ... → most famous/studied polaron models: Frohlich (continuous model) and Holstein (lattice model). → Today: review of methods to study the single-polaron problem in the Holstein model, + some results

Model of interest: the Holstein Hamiltonian (1959)

The simplest lattice Hamiltonian describing electron-phonon (phonons = lattice vibrations) interactions:

$$H = -t \sum_{\langle i,j \rangle,\sigma} (c_{i\sigma}^+ c_{j\sigma}^- + c_{j\sigma}^+ c_{i\sigma}^-) + \Omega \sum_i b_i^+ b_i^- + g \sum_i n_i (b_i^+ + b_i^-)$$
Kinetic energy – describes how the electron hops on the lattice

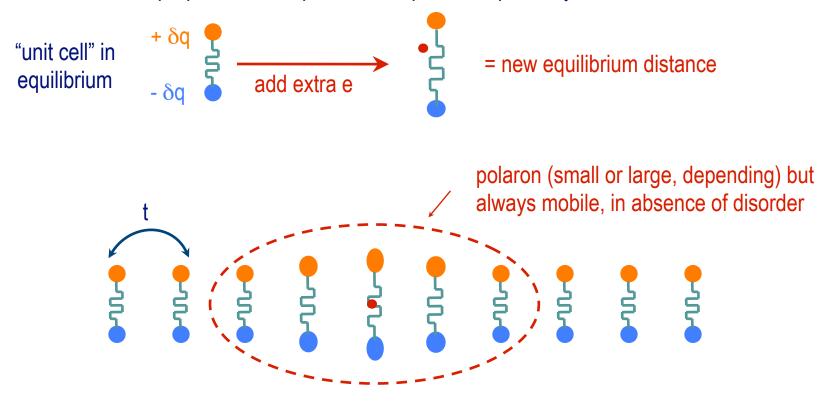
$$|attice|$$
Interaction: $n_i^- = \#$ of electrons at the site

Each atom of the lattice is like a harmonic oscillator (quick reminder):

$$h_{i} = \frac{\hat{P}_{i}^{2}}{2M} + \frac{M\Omega(\hat{X}_{i} - X_{0,i})^{2}}{2} \rightarrow \hbar\Omega\left(b_{i}^{+}b_{i} + \frac{1}{2}\right)$$

$$\begin{cases}
b_{i} = \sqrt{\frac{M\Omega}{2\hbar}} \left(\hat{X}_{i} - X_{0,i} + i\frac{\hat{P}_{i}}{M\Omega}\right) \\
b_{i}^{+} = \sqrt{\frac{M\Omega}{2\hbar}} \left(\hat{X}_{i} - X_{0,i} - i\frac{\hat{P}_{i}}{M\Omega}\right)
\end{cases} \rightarrow \hat{X}_{i} - X_{0,i} \propto (b_{i}^{+} + b_{i})$$

Hamiltonian was proposed as simplified description of a polar crystal → 1D sketch



Eigenstates are linear combinations of states with the electron at different sites, surrounded by a lattice distortion (cloud of phonons).

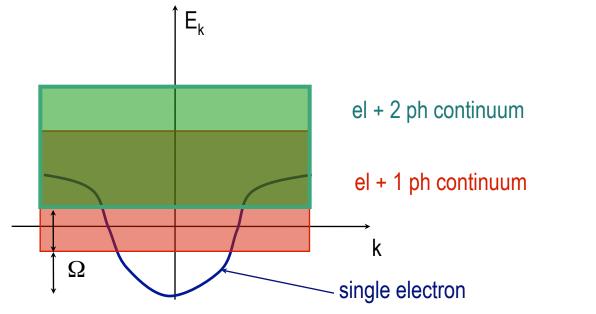
This composite object = electron dressed by surrounding cloud of phonons is called a polaron.

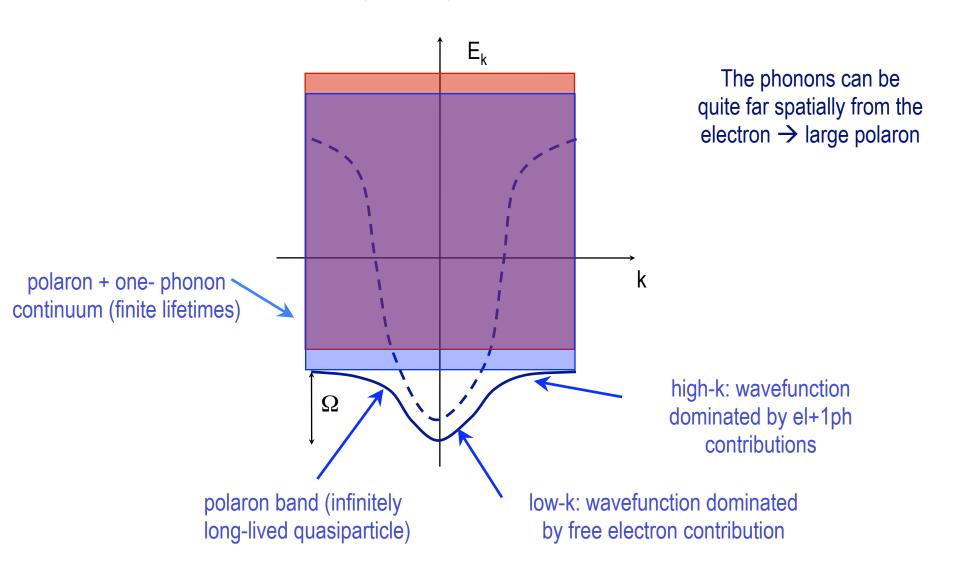
$$\begin{split} H &= -t \sum_{\langle i,j \rangle,\sigma} \left(c_{i\sigma}^{+} c_{j\sigma}^{} + c_{j\sigma}^{+} c_{i\sigma}^{} \right) + \Omega \sum_{i} b_{i}^{+} b_{i}^{} + g \sum_{i} n_{i}^{} (b_{i}^{+} + b_{i}^{}) \\ &= \sum_{k} \varepsilon_{\vec{k}} c_{\vec{k}}^{+} c_{\vec{k}}^{} + \Omega \sum_{\vec{q}} b_{\vec{q}}^{+} b_{\vec{q}}^{} + \frac{g}{\sqrt{N}} \sum_{k,\vec{q}} c_{\vec{k}-\vec{q}}^{+} c_{\vec{k}}^{} \left(b_{\vec{q}}^{+} + b_{-\vec{q}}^{} \right) \end{split}$$

(spin is irrelevant, N = number of unit cells, \rightarrow infinity at the end, all k,q-sums over Brillouin zone)

Asymptotic behavior:

ightharpoonup zero-coupling limit, g=0 ightharpoonup eigenstates of given k: $c_{\vec{k}}^+ \left| 0 \right\rangle, \ c_{\vec{k}-\vec{q}}^+ b_{\vec{q}}^+ \left| 0 \right\rangle, \ c_{\vec{k}-\vec{q}-\vec{q}}^+ b_{\vec{q}}^+ b_{\vec{q}}^+ \left| 0 \right\rangle, \dots$ with eigenenergies $\epsilon_{\vec{k}}, \ \epsilon_{\vec{k}-\vec{q}} + \Omega, \ \epsilon_{\vec{k}-\vec{q}-\vec{q}}^+ + 2\Omega, \dots$ where, for example, $\epsilon_{\vec{k}} = -2t \sum_{i=1}^d \cos k_i$





 \rightarrow infinitely strong coupling, t=0 \rightarrow electron stays at a single site forever \rightarrow n_i=1 there, 0 elsewhere

$$H = \left[\Omega b_0^{\dagger} b_0 + g \left(b_0^{\dagger} + b_0\right)\right] + \Omega \sum_{i \neq 0} b_i^{\dagger} b_i = \Omega B_0^{\dagger} B_0 - \frac{g^2}{\Omega} + \Omega \sum_{i \neq 0} b_i^{\dagger} b_i$$

where
$$B_0 = b_0 + \frac{g}{\Omega} \rightarrow \left[B_0, B_0^{\dagger} \right] = \left[b_0, b_0^{\dagger} \right] = 1$$

Ground-state is:
$$|GS\rangle = c_0^{\dagger} \left| -\frac{g}{\Omega} \right|_0 \rightarrow E_{GS} = -\frac{g^2}{\Omega}$$

where
$$b_0 \left| -\frac{g}{\Omega} \right|_0 = -\frac{g}{\Omega} \left| -\frac{g}{\Omega} \right|_0 \rightarrow B_0 \left| -\frac{g}{\Omega} \right|_0 = 0, b_i \left| -\frac{g}{\Omega} \right|_0 = 0$$

polaron binding energy; small polaron limit

$$N_{ph} = \left\langle b_0^+ b_0 \right\rangle_{GS} = \frac{g^2}{\Omega^2}$$

while excited states have energies $E = -\frac{g^2}{\Omega} + n\Omega$

only discrete eigenstates!

and eigenfunctions of the general form: $c_0^{\dagger} B_0^{\dagger^{m_0}} \prod_{i \neq 0} b_i^{\dagger^{m_i}} \left| -\frac{g}{\Omega} \right\rangle_0$, $n = \sum m_i$

Recall coherent states:
$$b|\alpha\rangle = \alpha |\alpha\rangle \otimes |\alpha\rangle = e^{-\frac{|\alpha|^2}{2} + \alpha b^{\dagger}} |0\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\left(b^{\dagger}\right)^n}{n!} |0\rangle$$

Ground-state energy: -2dt (half-bandwidth, at zero coupling) \rightarrow -g²/ Ω (infinite coupling)

 \rightarrow Effective coupling as their ratio $\lambda = g^2/(2dt\Omega) \rightarrow$ weak coupling $\lambda << 1$, strong coupling $\lambda >> 1$

3 energy scales: t, Ω , g \rightarrow 2 dimensionless parameters λ = $g^2/(2dt\Omega)$, Ω/t (d is lattice dimension)

 \rightarrow very strong coupling, $\lambda >> 1 \rightarrow$ polaron energy is

$$E_k = -\frac{g^2}{\Omega} + e^{-\frac{g^2}{\Omega^2}} \varepsilon_k + \dots \rightarrow t_{eff} = te^{-\frac{g^2}{\Omega^2}} \rightarrow m_{eff} = me^{\frac{g^2}{\Omega^2}}$$

and wavefunction is
$$|\psi_k\rangle = \sum_i \frac{e^{i\vec{k} \cdot \vec{R_i}}}{\sqrt{N}} c_i^{\dagger} \left| -\frac{g}{\Omega} \right\rangle_i$$

Again, must have a polaron+one-phonon continuum at $E_{GS} + \Omega \rightarrow$ details too nasty

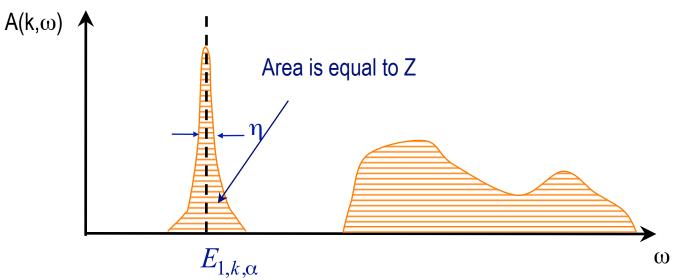
Question: how is the spectrum evolving between these two very different limits?

Quantity of interest: the Green's function or propagator

$$H | 1, k, \alpha \rangle = E_{1,k,\alpha} | 1, k, \alpha \rangle$$
 \leftarrow eigenenergies and eigenfunctions (1 electron, total momentum k, α is collection of other needed quantum numbers)

$$G(k,\omega) \triangleq \left\langle 0 \left| c_k \frac{1}{\omega - H + \dot{\eta}} c_k^+ \left| 0 \right\rangle = \sum_{\alpha} \frac{Z_{1,k,\alpha}}{\omega - E_{1,k,\alpha} + \dot{\eta}} \qquad Z_{1,k,\alpha} = \left| \left\langle 1, k, \alpha \left| c_k^+ \left| 0 \right\rangle \right|^2 \right| \right\rangle$$

$$A(k,\omega) \triangleq -\frac{1}{\pi} \operatorname{Im} G(k,\omega)$$
 \leftarrow = spectral weight, is measured (inverse) angleresolved photoemission spectroscopy (ARPES)



Z = quasiparticle weight → measures how similar is the true wavefunction to a non-interacting (free electron, no phonons) wavefunction

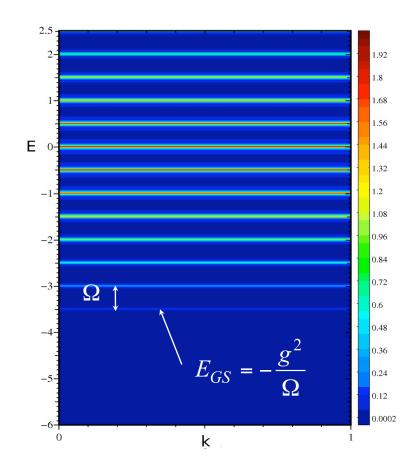
weak coupling
$$\lambda = \frac{g^2}{2dt\Omega} = 0$$
 $(g = 0)$

$$G_0(k,\omega) = \frac{1}{\omega - \varepsilon_k + i\eta};$$

$$A_0(k,\omega) = \frac{\eta}{\pi \left[\left(\omega - \varepsilon_k \right)^2 + \eta^2 \right]} \xrightarrow{\eta \otimes 0} \delta \left(\omega - \varepsilon_k \right)$$

Lang-Firsov impurity limit
$$\lambda = \frac{g^2}{2dt\Omega} = \infty$$
 $(t = 0)$

$$G_{LF}(k,\omega) = e^{-\frac{g^2}{\Omega^2}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{g}{\Omega}\right)^{2n} \frac{1}{\omega + \frac{g^2}{\Omega} - n\Omega + i\eta}$$



How does the spectral weight evolve between these two very different looking limits?

Most numerical approaches → focus on the evolution of the polaron band (low-energy properties)

variational methods (Trugman and co-workers)

$$\left| \psi_{k} \right\rangle = \sum_{i} e^{i\vec{k} \times \vec{R}_{i}} c_{i}^{\dagger} \left[1 + \sum_{\delta} \varphi_{\delta} b_{i+\delta}^{\dagger} + \sum_{\delta,\delta'} \varphi_{\delta,\delta'} b_{i+\delta}^{\dagger} b_{i+\delta'}^{\dagger} + \dots \right] \left| 0 \right\rangle$$

- → truncate size of cloud (both spatial and how many phonons are allowed) → Lanczos
- → advantages: continuous k (not a finite-size chain!); matrix elements are very simple to get, can be extremely accurate for discrete states (like the polaron band of interest)
- \rightarrow disadvantages: at large couplings, very many phonon combinations \rightarrow huge dimension of variational Hilbert space (gets worse in higher dimension). Also, no predictive powers for the continuum above the polaron band \rightarrow nothing about high-energy properties.

- Diagrammatic Quantum Monte Carlo (Prokof'ev, Svistunov and co-workers)
- → calculate Green's function in imaginary time

$$G(k,\tau) = \langle 0 | c_k e^{-\tau H} c_k^{\dagger} | 0 \rangle = \sum_{\alpha} e^{-\tau E_{1,k,\alpha}} \left| \langle 1, k, \alpha | c_k^{\dagger} | 0 \rangle \right|^2 \xrightarrow{\tau \to \infty} Z_k e^{-\tau E_k}$$

Basically, use Metropolis algorithm to sample which diagrams to sum, and keep summing numerically until convergence is reached

- \rightarrow advantages: once code is written, it is fast (min. per data point) and very accurate for low-energies (discrete eigenstates). In principle it can be used to generate whole G(k,w) but convergence for short-times is much more difficult, also one needs analytic continuation to switch to real frequencies \rightarrow A. Mishchenko
- → disadvantages: writing the code (for me, at least)

- ➤ Quantum Monte Carlo methods (Kornilovitch in Alexandrov group, Hohenadler in Fehske group, ...) → write partition function as path integral, use Trotter to discretize it, then evaluate. Mostly low-energy properties are calculated/shown.
- ightharpoonup Exact diagonalization = ED ightharpoonup finite system (still need to truncate Hilbert space) ightharpoonup can get whole spectrum and then build G(k,w)
- ➤ Cluster perturbation theory: ED finite system, then use perturbation in hopping to "sew" finite pieces together → infinite system.
- \rightarrow advantage: can calculate G(k,w) for all k.
- → disadvantage: problems in higher dimension, and lower couplings (big phonon clouds)
- ➤ "Special" methods:
- → DMRG (density matrix renormalization group, if d=1)
- → DMFT (dynamic mean-field theory, if d→ infinity)
- > (lots of work done in these 50 years, as you may imagine)

They're all in very good agreement for low-energy properties, the difference is in efficiency and "generalizability" to higher dimensions, other models, etc.

Analytic approaches (other than perturbation theory) → calculate self-energy

$$G(k,\omega) = \frac{1}{\omega - \varepsilon_k - \Sigma(k,\omega) + i\eta}$$

$$\Sigma(k,\omega) = \frac{1}{2} + \frac{1}{$$

For Holstein polaron, we need to sum to orders well above g^2/Ω^2 to get convergence.

n	1	2	3	4	5	6	7	8
Σ , exact	1	2	10	74	706	8162	110410	1708394
Σ, SCBA	1	1	2	5	14	42	132	429

Traditional approach: find a subclass of diagrams that can be summed, ignore the rest

→ self-consistent Born approximation (SCBA) – sums only non-crossed diagrams (much fewer)

New proposal: the $MA^{(n)}$ hierarchy of approximations:

Idea: keep ALL self-energy diagrams, but approximate each such that the summation can be carried out analytically. (Alternative explanation: generate the infinite hierarchy of coupled equations of motion for the propagator, keep all of them instead of factorizing and truncating, but simplify coefficients so that an analytical solution can be found).

First: MA⁽⁰⁾ – simplest (least accurate) version

Replace each ——— in the self-energy diagrams by $G_0(\vec{k},\omega)$

→ one can sum all the resulting self-energy diagrams:

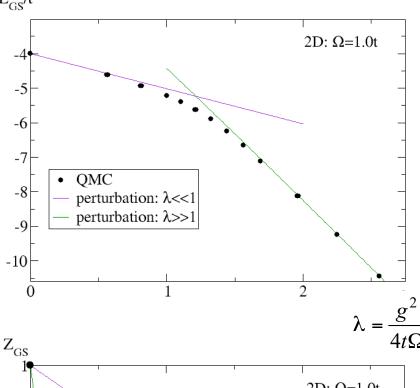
$$\sum_{MA^{(0)}}(\omega) = \frac{g^2 g_0(\omega - \Omega)}{1 - \frac{2g^2 g_0(\omega - \Omega)g_0(\omega - 2\Omega)}{1 - \frac{3g^2 g_0(\omega - 2\Omega)g_0(\omega - 3\Omega)}}$$
 \(\sim \text{trivial to evaluate}

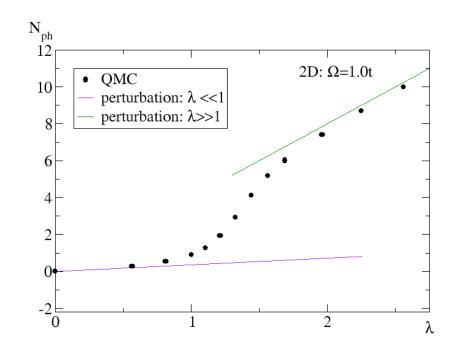
$$g_0(\omega) = \frac{1}{N} \sum_{k} G_0(\vec{k}, \omega)$$

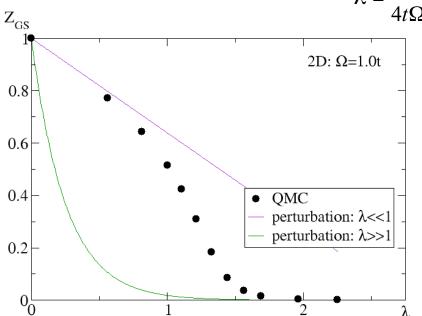
$$= \int_{B.Z.} \frac{d\vec{k}}{(2\pi)^d} \frac{1}{\omega - \varepsilon_{\vec{k}} + i\eta}$$

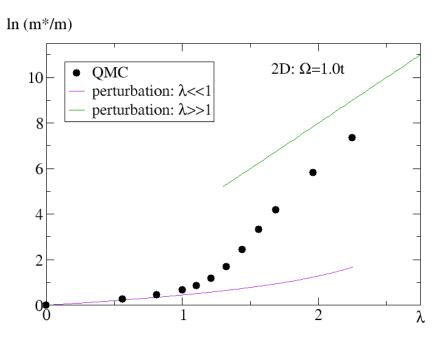
- ← result is EXACT both for g=0 and for t=0
- → There are good reasons why this should work well at low energies (ask!)
- → This approx. obeys exactly multiple sum rules for the spectral weight (at least 6)

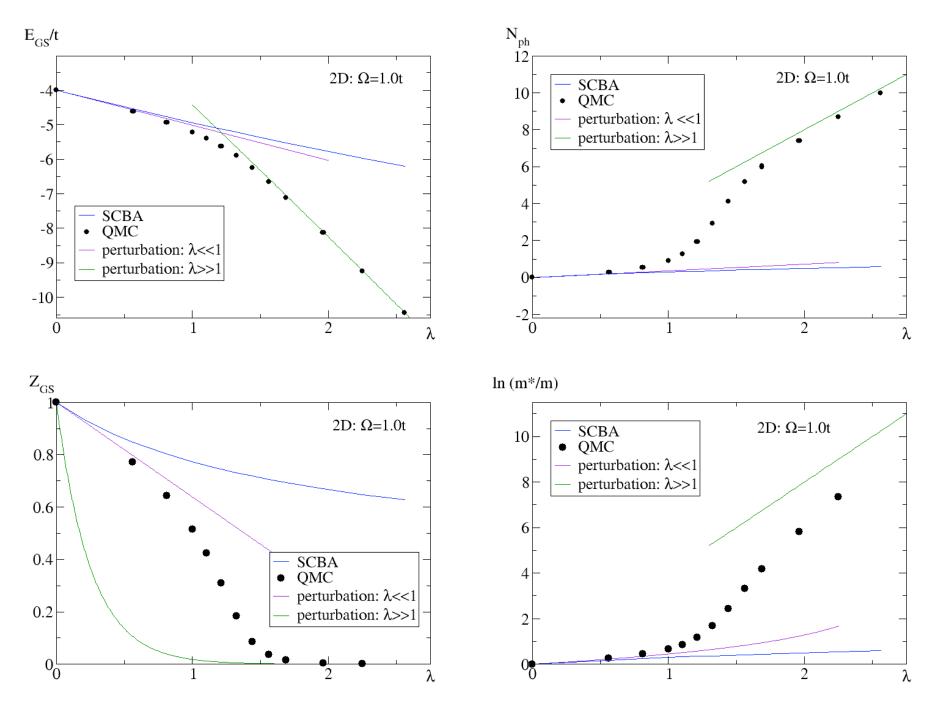
${}^{\mbox{2D}}_{\rm E_{\rm GS}/t}$ results for ground-state properties

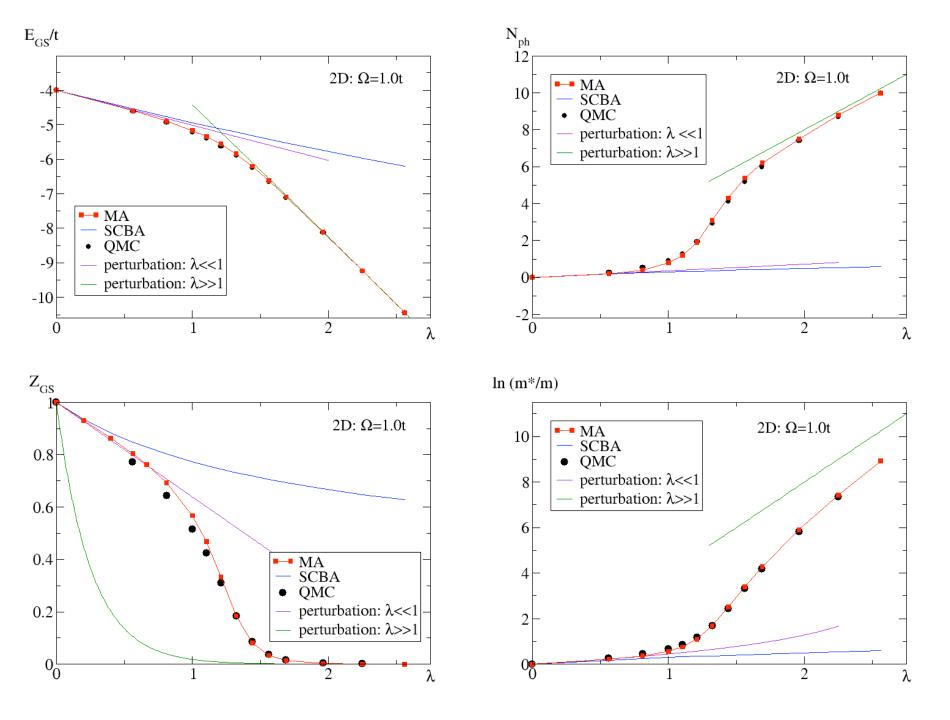




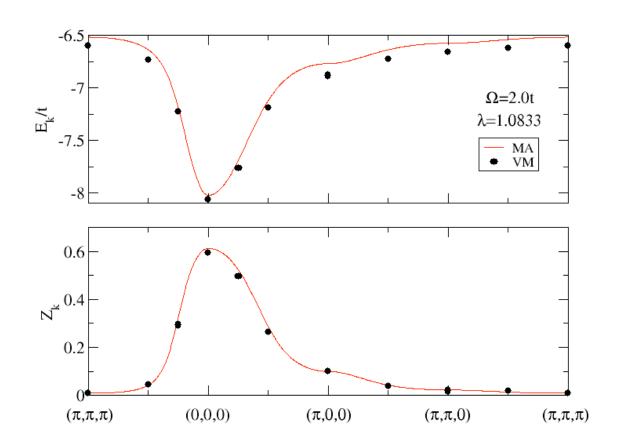




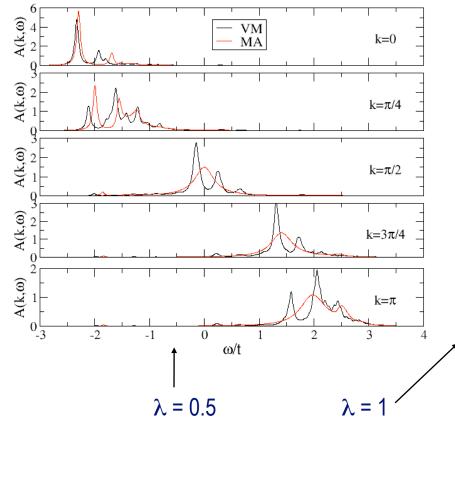




3D Polaron dispersion



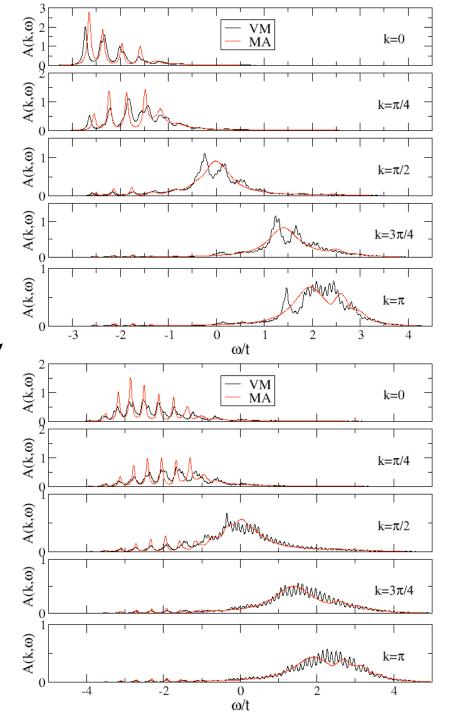
L. -C. Ku, S. A. Trugman and S. Bonca, Phys. Rev. B 65, 174306 (2002).



A(k, ω) in 1D, Ω =0.4 t

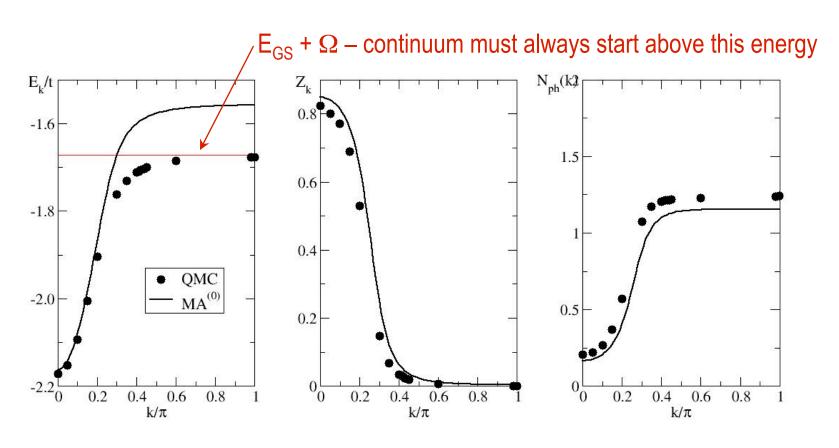
G. De Filippis et al, PRB 72, 014307 (2005)

MA becomes exact for small, large λ



MA⁽⁰⁾ is remarkably good, especially considering how simple it is. Higher d is equally trivial as d=1. However, it is an approximation, and it does have its problems:

- → self-energy is momentum independent
- \rightarrow the accuracy worsens if $\Omega/t \rightarrow 0$
- → doesn't see the polaron+one phonon continuum where it should be



1D, Ω =0.5t, λ =0.25

Problems easy to fix \rightarrow improve the approximation:

 $\mathsf{MA}^{(n)}$ keep free propagators of frequency $\omega - \mathsf{m}\Omega$, $\mathsf{m} < \mathsf{n}$ exactly in the self-energy diagrams; all propagators with more phonons (lower energy) are momentum averaged

 $MA^{(1)} - G_0(k-q,\omega-\Omega)$ contributions exact, lines with 2 or more phonons are momentum averaged.

 $MA^{(2)} - G_0(k-q,\omega-\Omega)$, $G_0(k-q,\omega-2\Omega)$ contributions exact, lines with 3 or more phonons are momentum averaged, etc.

Still can sum all diagrams in the self-energy, calculation still numerically trivial

Define continued fractions:
$$A_n(\omega) = \frac{ng_0(\omega - n\Omega)}{1 - g^2g_0(\omega - n\Omega)A_{n+1}(\omega)}$$

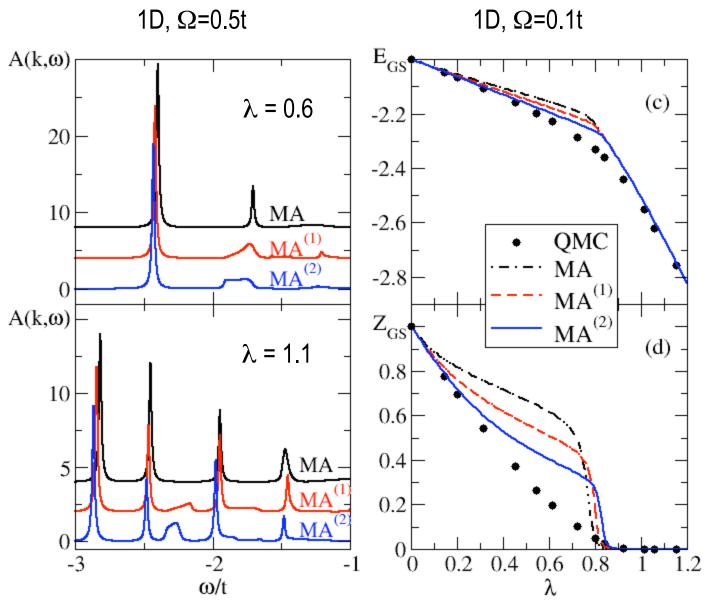
$$\Sigma_{MA^{(0)}}(\omega) = g^2 A_1(\omega)$$

$$\Sigma_{MA^{(1)}}(\omega) = \frac{g^2 g_0(\omega - \Omega - g^2 A_1(\omega - \Omega))}{1 - g^2 g_0(\omega - \Omega - g^2 A_1(\omega - \Omega)) \left[A_2(\omega) - A_1(\omega - \Omega)\right]}$$

 $\Sigma_{MA^{(2)}}(k,\omega) = \dots \rightarrow$ acquires explicit momentum dependence

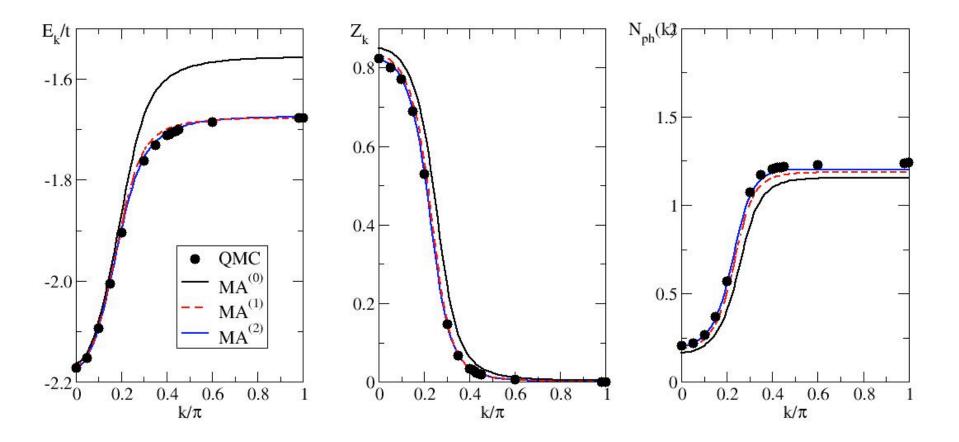
details in PRB 76, 165109 (2007)

(models with g(q) coupling have a k-dependent self-energy from level $MA^{(0)}$)

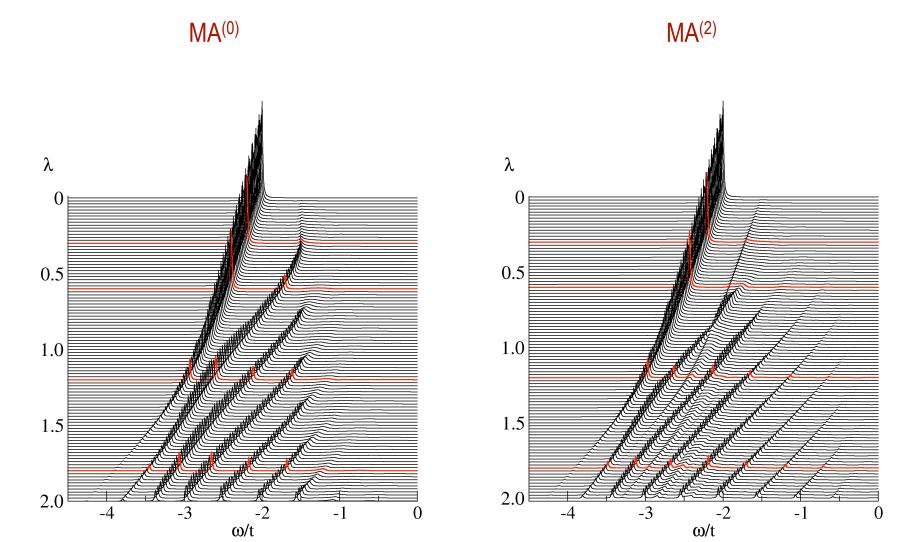


Sum rules:

MA⁽⁰⁾ exact up to n=5 and accurate above; MA⁽¹⁾ exact up to n=7 and more accurate above; MA⁽²⁾ exact up to n=9 and yet more accurate above, ...

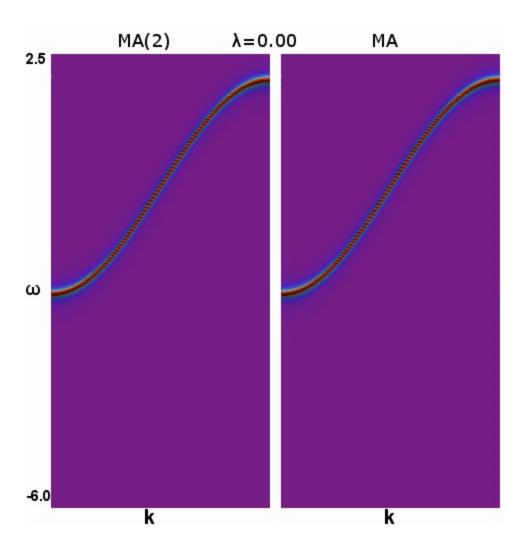


1D, Ω =0.5t, λ =0.25



1D, k=0, $\Omega=0.5t$

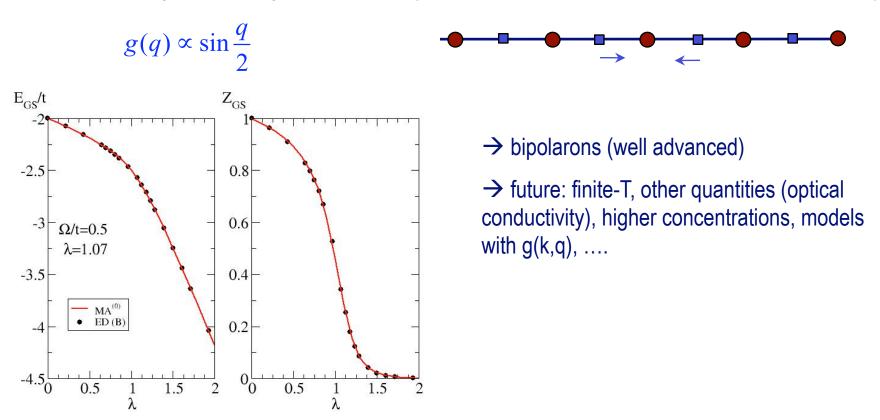
Our answer to how spectral weight evolves as λ increases from weak to strong coupling



Generalizations:

- → Lucian: multiple phonon modes and/or multiple free-electron bands (but Holstein coupling)
- → Glen: el-ph coupling which depends on phonon momentum

Example: coupling to breathing-mode phonon (phonons live on different sublattice than the electron)



Numerics: Bayo Lau, M. Berciu and G. A. Sawatzky, PRB 76, 174305 (2007)

Why should this be a reasonable thing to do?

(i) Real-space argument: $\mathsf{MA}^{(0)}$ means $G_0(i-j,\omega-n\Omega) \to \delta_{i,j}G_0(0,\omega-n\Omega) = \delta_{i,j}g_0(\omega-n\Omega)$

$$i \qquad \qquad exact: \quad g^4 \sum_{i,j} G_0(j-i,\omega-\Omega)G_0(i-j,\omega-2\Omega)G_0(j-i,\omega-\Omega)$$

$$j \qquad \qquad MA^{(0)}: \quad g^4 g_0(\omega-\Omega)g_0(\omega-2\Omega)g_0(\omega-\Omega)$$

At low energies $\omega \sim E_{GS} < -2dt \rightarrow$ free electron Greens' functions decrease *exponentially* with distance $|i-j| \rightarrow MA^{(0)}$ keeps the most important (diagonal) contribution. The approximation becomes better the more phonons are present, since the lower $\omega - n \Omega$ is, the faster the decay.

→ Expect ground-state properties to be described quite accurately.

(ii) Spectral weight sum rules (see PRB 74, 245104 (2006) for details)

(ii) Spectral weight sufficiency (see PRB 74, 245104 (2006) for details)
$$M_n(k) = \int_{-\infty}^{\infty} d\omega \omega^n A(k, \omega) = -\frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{\infty} d\omega \omega^n G(k, \omega) \qquad \leftarrow \text{ can be calculated exactly}$$
$$M_n(k) = \left\langle 0 \middle| c_k H^n c_k^+ \middle| 0 \right\rangle$$

MA⁽⁰⁾ satisfies exactly the first 6 sum rules, and with good accuracy all the higher ones.

Note: it is not enough to only satisfy a few sum rules, even if exactly. ALL must be satisfied as well as possible.

Examples: 1. SCBA satisfies exactly the first 4 sum rules, but is very wrong for higher order sum rules \rightarrow fails miserably to predict strong coupling behavior (proof coming up in a minute).

2. Compare these two spectral weights:

$$A_1(\omega) = \delta(\omega) \otimes M_0 = 1; M_{n>0} = 0$$

$$-\mathbf{w}_0 \qquad \mathbf{0} \qquad \mathbf{w}_0$$

$$A_2(\omega) = \frac{1}{2} \left(\delta(\omega - \omega_0) + \delta(\omega + \omega_0) \right) \otimes M_n = \frac{\omega_0^n}{2} \left[1 + (-1)^n \right] = 0, \text{ if n is odd}$$

$$M_n(k) = \int_{-\infty}^{\infty} d\omega \omega^n A(k, \omega) = -\frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{\infty} d\omega \omega^n G(k, \omega)$$

Since G(k,w) is a sum of diagrams, keeping the correct no. of diagrams is extremely important!

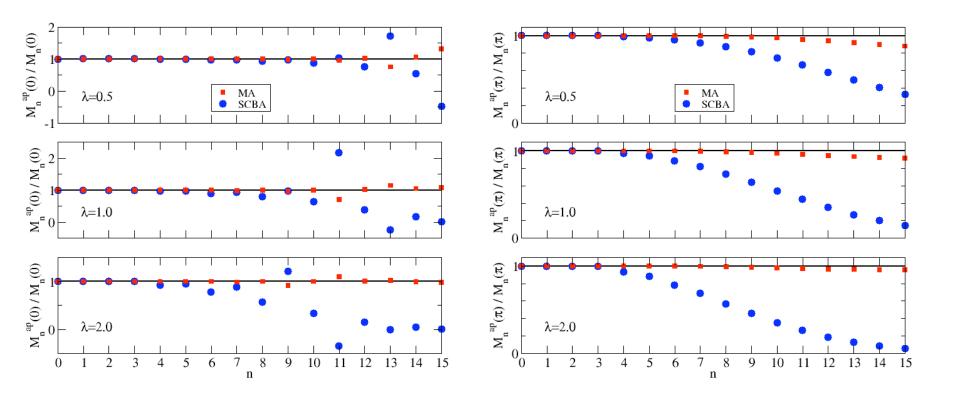
found correctly if n=0 diagram kept correctly \rightarrow dominates if t >> g, $\lambda \rightarrow$ 0

$$M_{6}(\vec{k}) = \varepsilon_{\vec{k}}^{6} + g^{2} \left[5\varepsilon_{\vec{k}}^{4} + 6t^{2} \left(2d^{2} - d \right) + 4\varepsilon_{\vec{k}}^{3} \Omega + 3\varepsilon_{\vec{k}}^{2} \Omega^{2} + 6dt^{2} \left(\varepsilon_{\vec{k}}^{2} + \varepsilon_{\vec{k}} \Omega + 2\Omega^{2} \right) + 2\varepsilon_{\vec{k}} \Omega^{3} + \Omega^{4} \right] + g^{4} \left[18dt^{2} + 12\varepsilon_{\vec{k}}^{2} + 22\varepsilon_{\vec{k}} \Omega + 25\Omega^{2} \right] + 15g^{6}$$

$$M_{6,MA}(\vec{k}) = M_6(\vec{k}) - 2dt^2g^4$$

found correctly if we sum correct no. of diagrams \rightarrow dominates if g >>t, λ >>1

$$M_{6,SCBA}(\vec{k}) = M_6(\vec{k}) - g^4 [....] - 10g^6$$



$$\lambda = \frac{g^2}{2dt\Omega}$$