Study of Polaron Properties using the Momentum Average Approximation

by

Glen L. Goodvin

B.Sc., The University of Northern British Columbia, 2003
M.Sc., The University of Northern British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
September 2009
© Glen L. Goodvin 2009
Abstract

In this work we present a highly efficient and accurate analytical approximation for the Green's function of a polaron: the momentum average (MA) approximation. It is obtained by summing all of the self-energy diagrams, but with each diagram averaged over the momenta of its free propagators. The result becomes exact for both zero bandwidth and for zero electron-phonon coupling, and is accurate everywhere in the parameter space.

The approximation is first used to investigate the Holstein model. A detailed analysis of its accuracy is provided through diagrammatics and spectral weight sum rules. It is shown that the resulting Green's function satisfies exactly the first six spectral weight sum rules, and all higher order sum rules are satisfied with great accuracy. Comparison with numerical data also confirms this accuracy. We then show how to improve the MA approximation by systematically improving the accuracy of the self-energy diagrams in such a way that they can still all be summed efficiently. This allows us to fix some of the problems of the zeroth-order MA approximation. The quantitative agreement with numerical data is also improved.

Next, we generalize the MA approximation to study the properties of models with momentum-dependent electron-phonon coupling, and then show that further improvements can be obtained based on variational considerations, using the 1D breathing-mode Hamiltonian as a specific example. For example, by using this variational MA, we obtain ground state energies within at most 0.3% error of the numerical data.

Finally, we study the effects of a nearby surface on the spectral weight of a Holstein polaron. The broken translational symmetry is accounted for without any additional approximations, and the resulting inhomogeneous MA approximation continues to be accurate for all coupling strengths. We show that the surface changes properties significantly, with bulk values being recovered only very far away from it. We find that the electron-phonon coupling gives rise to an additional surface potential which is responsible for binding surface states even when they are not normally expected. These results demonstrate that interpretation in terms of bulk properties of spectroscopic data sensitive only to a few surface layers is not straightforward.
# Table of Contents

Abstract .................................................................................. ii

Table of Contents ................................................................. iii

List of Figures ................................................................. vi

Acknowledgements ......................................................... viii

Statement of Co-Authorship ................................................ ix

1 Introduction ................................................................. 1

1.1 The Electron-Phonon Interaction .......................... 1

1.2 Model of the Electron-Phonon Interaction .......... 2

1.3 The Green’s Function ............................................. 5

1.4 The Holstein Model ................................................. 6

1.4.1 Analytical Techniques ................................. 7

1.4.2 Numerical Methods .................................. 12

1.5 The Breathing-mode Hamiltonian ................. 12

1.5.1 Analytical Techniques ................................. 13

1.6 The Highly Inhomogeneous Coupling Model .... 16

1.7 Experimental Evidence for Polarons .......... 17

1.8 Motivation .......................................................... 20

2 The Holstein Polaron .................................................. 22

2.1 The Momentum Average Approximation ......... 23

2.1.1 The Exact Solution ................................... 23

2.1.2 The Approximation ................................... 26

2.2 The Meaning of the MA Approximation ......... 28

2.2.1 Diagrammatics ........................................ 28

2.2.2 Variational Interpretation of the MA Approximation 30

2.2.3 Sum Rules for the MA Approximation ......... 33

2.3 Results ............................................................... 39
# Table of Contents

2.3.1 Polaron Ground State Properties ........................................ 40  
2.3.2 Low Energy States: Momentum Dependence .......................... 49  
2.3.3 Higher Energy States .................................................. 51  
2.4 Conclusion ..................................................................... 70

3 Systematic Improvement of the MA Approximation .................. 72  
3.1 Higher Level MA .............................................................. 73  
3.1.1 MA(1) Level .................................................................. 74  
3.1.2 MA(2) Level .................................................................. 76  
3.2 Results ........................................................................... 80  
3.2.1 Ground State Properties ............................................... 80  
3.2.2 Polaron Band ................................................................ 82  
3.2.3 Higher Energy States .................................................... 82  
3.3 Conclusion ..................................................................... 96

4 Generalization to Models with Electron-Phonon Coupling Dependent on the Phonon Momentum ......................... 97  
4.1 The Model ....................................................................... 98  
4.2 Calculating the Green’s Function ........................................ 98  
4.2.1 Exact Solution ............................................................ 98  
4.2.2 The MA(0) Approximation ......................................... 100  
4.2.3 The MA(1) Approximation ......................................... 102  
4.2.4 MA(0) and MA(1) Results .......................................... 103  
4.3 The Variational MA Approximation .................................... 107  
4.3.1 The MA(v,0) Approximation ...................................... 107  
4.3.2 The MA(v,1) Approximation ...................................... 110  
4.4 Variational MA Results .................................................... 112  
4.4.1 Ground State Properties ............................................. 112  
4.4.2 The Polaron Band ....................................................... 113  
4.4.3 Higher Energy Properties .......................................... 116  
4.5 Conclusion ..................................................................... 119

5 The Effect of a Surface on a Holstein Polaron ....................... 122  
5.1 The Model of the Surface ................................................ 123  
5.2 The Holstein Model with a Surface ................................... 127  
5.2.1 Exact Solution ............................................................ 127  
5.2.2 The IMA Approximation .......................................... 129  
5.3 IMA Results .................................................................. 133  
5.3.1 Isotropic Hopping ..................................................... 133  
5.3.2 Anisotropic Hopping .................................................. 139
## Table of Contents

5.4 Conclusion ................................................. 142

6 Conclusion ................................................. 145  
   6.1 Summary of this Work ................................. 145  
   6.2 Further Applications of MA ......................... 148

Bibliography .................................................. 149

Appendices .................................................... 154  
   A Comparison with DMFT ................................. 154  
   B MA\(^{(1)}\) Details ...................................... 156  
   C MA\(^{(2)}\) Details ...................................... 158  
   D Momentum Averages ..................................... 162  
   E MA\(^{(v,1)}\) Details .................................... 164  
   F IMA\(^{(1)}\) Details ..................................... 167  
   G List of Publications ................................... 170
## List of Figures

1.1 Number of diagrams included in the SCBA self-energy . . . . 10
1.2 The quasiparticle kink in high-$T_c$ from Lanzara et al. . . . 17
1.3 Evidence for polarons in high-$T_c$ from Shen et al. . . . . 19
1.4 MA fit to the spectral data of Shen et al. . . . . . . . . . . 20

2.1 Diagrammatical expansion for the self energy . . . . . . . . 25
2.2 Second order Feynman diagram in real space . . . . . . . . 31
2.3 Ratio of MA, SCBA sum rules to exact sum rules, $k = 0$ . . 36
2.4 Ratio of MA, SCBA sum rules to exact sum rules, $k = \pi$ . 37
2.5 1D ground state results for $E_0$ and $Z_0$ . . . . . . . . . . 42
2.6 1D ground state results for $N_{ph}$ and $m^*$ . . . . . . . . 43
2.7 2D ground state results for $E_0$ and $Z_0$ . . . . . . . . . . 45
2.8 2D ground state results for $N_{ph}$ and $m^*$ . . . . . . . . . 46
2.9 3D ground state results for $E_0$ and $Z_0$ . . . . . . . . . . 47
2.10 3D ground state results for $N_{ph}$ and $m^*$ . . . . . . . . . 48
2.11 Expectation values for components of the polaron energy . 50
2.12 Lowest polaron eigenenergy and $qp$ weight versus $k$ in 1D . 52
2.13 Number of phonons in the polaronic cloud versus $k$ in 1D . 53
2.14 Ground state energy and $qp$ weight versus $k$ in 2D . . . . . 54
2.15 Number of phonons in the polaronic cloud versus $k$ in 2D . 55
2.16 Ground state energy and $qp$ weight versus $k$ in 3D . . . . . 56
2.17 Higher energy properties in 1D . . . . . . . . . . . . . . . . 58
2.18 Comparison of $A(k, \omega)$ to variational results with $\lambda = 0.5$ . . 59
2.19 Comparison of $A(k, \omega)$ to variational results with $\lambda = 1.0$ . 60
2.20 Comparison of $A(k, \omega)$ to variational results with $\lambda = 2.0$ . 61
2.21 The spectral weight for varying $\eta$ in 1D . . . . . . . . . . 62
2.22 Comparisons of $A(k, \omega)$ with ED results in 2D . . . . . . 63
2.23 $qp$ weight and $N_{ph}$ for higher order bound states . . . . . 64
2.24 Spectral weight $A(k, \omega)$ versus $k$ and $\omega$ in 1D . . . . . 66
2.25 $A(k, \omega)$ and $A(\omega)$ versus $\lambda$ and $\omega$ in 1D . . . . . 67
2.26 $A(k, \omega)$ and $A(\omega)$ versus $\lambda$ and $\omega$ in 2D and 3D . . . . . 68
### List of Figures

3.1 1D results for $E_0$ and $Z_0$ using higher order MA  
3.2 1D results for $E_k$ and $Z_k$ using higher order MA  
3.3 1D results for $N_{ph}(k)$ using higher order MA  
3.4 1D spectral results using higher order MA, $\lambda = 0.6$  
3.5 1D spectral results using higher order MA, $\lambda = 1.2$  
3.6 1D spectral results using higher order MA, $\lambda = 1.8$  
3.7 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA  
3.8 Appearance of a 3rd bound state below the continuum in 3D  
3.9 $N_{ph}(k)$ using higher order MA  
3.10 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA, $\lambda = 0.6$  
3.11 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA, $\lambda = 1.2$  
3.12 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA, $\lambda = 1.8$  
3.13 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA, $\lambda = 1.95$  
3.14 $A(k,\omega)$ vs. $\lambda$ and $\omega$ using higher order MA, $\lambda = 1.2$  
4.1 $E_0$ and $Z_0$ for the 1D breathing-mode, using MA$^{(n)}$  
4.2 Comparison to HIC data using MA$^{(n)}$  
4.3 Pictorial representation generalized Green’s functions  
4.4 Ground state results for the breathing-mode, using MA$^{(v,n)}$  
4.5 $N_{ph}$ and $m^*$ for the 1D breathing-mode, using MA$^{(v,n)}$  
4.6 $E_k$ and $Z_k$ for the 1D breathing-mode, using MA$^{(v,n)}$  
4.7 $A(0,\omega)$ vs. $\omega$ for several couplings, using MA$^{(v,n)}$  
4.8 $A(k,\omega)$ vs. $\omega$ for several momenta, using MA$^{(v,n)}$  
4.9 $A(0,\omega)$ vs. $\lambda$ and $\omega$, using MA$^{(v,n)}$  
5.1 $A(k,\omega)$ vs. $\omega$, with $g = 0$  
5.2 $A(k,\omega)$ vs. $\omega$, with $U = 0$, using IMA$^{(n)}$  
5.3 $A(k,\omega)$ vs. $\omega$, with $U = 0$, using IMA$^{(n)}$  
5.4 $A(k,\omega)$ vs. $\omega$, with $U/t = 0.5$, using IMA$^{(n)}$  
5.5 Effect of $v_1$($\omega$) on $A(k=0,i,i,\omega)$, for $U/t = 0.5$  
5.6 Energy and qP weight of the lowest bound state vs. $U/t$  
5.7 Components of the effective potential for $U = 0$  
5.8 $A(k,i,i,\omega)$ vs. $k$ and $\omega$, with $U/t = 0.5$, using IMA$^{(l)}$  
5.9 $A(k=0,i,i,\omega)$ vs. $\omega$ and $t/t||$  
5.10 $A(k=0,i,i,\omega)$ vs. $\omega$ and $t/t||$, with $U/t = 0.5$  
5.11 Energy and qP weight of the lowest bound state vs. $U/t||$  
A.1 Comparison between MA and DMFT  

vii
Acknowledgements

Above all, I would like to thank my supervisor Mona Berciu. Her unwa-
vering support and guidance throughout my studies have always provided
an excellent learning and working environment, and her rare combination of
extraordinary skill, modesty, and kindness have made it a true pleasure and
inspiration to work with her over the last four years.

During my stay at the University of British Columbia I have benefited
greatly from relationships with my colleagues and friends. I would especially
like to thank Lucian Covaci, Roman Holenstein, Justin Malecki, Dominic
Marchand, Blair Tennessy, Conan Weeks, and Mike & Amy Wood. I am also
very thankful to have had the support and guidance of my M.Sc. supervisor
and friend, Prof. Mark R.A. Shegelski.

I would also like to thank my Ph.D. committee members, Prof. Andrea
Damascelli, Dr. Marcel Franz, and Prof. George Sawatzky for their advice
and constructive feedback over the course of my studies. In addition I would
like to explicitly acknowledge the co-authors of the publications resulting
from this work: Dr. Mona Berciu, Dr. Lucian Covaci, and Prof. George
Sawatzky.

On a personal level, I am very thankful of my girlfriend Caitlin Mischki,
for her company, for standing behind me, for pushing me to do my best,
and for her support during the final stretch. Finally, I would like to thank
my parents. I am extremely grateful for your encouragement throughout
my life, and I am forever indebted to you both for always providing me with
the opportunity to achieve my goals.
Statement of Co-Authorship

This thesis is based on a series of publications authored by me, my supervisor Mona Berciu (Chapters 2, 3, 4, 5), and Lucian Covaci (Chapter 5). These publications are listed in Appendix G. The content of the thesis is based on the original work of Mona Berciu, published in Physical Review Letters 97, 036402 (2006). My research program followed naturally from this pioneering work, with both me and my supervisor involved in formulating the research topics leading to the publications above. The research and manuscript preparation were roughly shared between me and my supervisor for Chapters 2 and 3, while the data collection was primarily my responsibility. I was responsible for performing the majority of the research and data collection for Chapters 4 and 5. The preparation of the manuscript corresponding to Chapter 4 was largely my responsibility, and the manuscript on which Chapter 5 is based was prepared by me and my supervisor.
Chapter 1

Introduction

1.1 The Electron-Phonon Interaction

There is a considerable interest in understanding the coupling of a particle to its environment, and in particular, the coupling between a charged particle and a lattice. The problem investigated in this thesis is that of an electron coupled to lattice vibrations. The resulting quasiparticle, an electron dressed by a phonon cloud, is called a polaron, a concept introduced by Landau many years ago [1]. The physics of polarons is relevant in many physical systems, including polymers, nanotubes, C_{60} and other fullerenes [2–4], manganites [5], Bechgaard salts [6, 7], and possibly in high-T_{c} compounds [8–11]. The lattice vibrations can not only dress the charge carrier to form polarons, but they can also induce an attractive interaction between two electrons. As Cooper demonstrated, a Fermi liquid is unstable to an attractive interaction between two electrons, and the resulting Cooper instability is a key ingredient in the formation of Cooper pairs and the development of the Bardeen, Cooper, and Schrieffer (BCS) theory of conventional superconductivity [12].

An important feature of a polaron is its increased mass, arising from the el-ph induced distortion carried by a dressed electron as it moves through a crystal. This quasiparticle is called a “large polaron” when the spatial extent of the lattice distortion is large compared to the lattice parameter, and it is called a “small polaron” when the lattice distortion is localized to within a few sites of the electron. The microscopic structure of these two extreme cases is very different, and it is correlated strongly with the strength and nature of the el-ph interaction. A large polaron is usually formed at small couplings, or for very long range couplings (or some combination of both), and it is signified by a slightly increased effective mass and a large-range but small lattice distortion. On the other hand, the small polaron has an exponentially large effective mass and is strongly localized by large lattice distortions within a few lattice sites of the electron. This occurs when the el-ph coupling is large and possesses short-range coupling.

The evolution from the large polaron to the small polaron as the el-ph
coupling is increased is in itself a very interesting problem. In the former
case the electron is lightly dressed and is described by wide-band Bloch
states, while in the latter case the electron is strongly localized is described
by a sequence of flat and narrow bands separated by the phonon excitation
energy. For the ground state it has been proven that the evolution between
these two very different extremes is analytic, and is therefore not described
by a phase transition [13]. However, there is still a sharp crossover from
the large to small polaron, and the physics in the intermediate regime is
both interesting and impossible to treat using existing analytical methods.
It is one of our main goals to better understand this crossover using the
momentum average (MA) approximation developed in this work.

The single electron problem that we address in this thesis is relevant
for insulators and possibly for weakly doped systems with low electron con-
centrations. For metals the effects of electron-electron interactions, direct
and/or mediated by phonons, must be considered. We do not address finite
electron concentrations in this work.

1.2 Model of the Electron-Phonon Interaction

It is instructive to derive a general Hamiltonian for describing electron-
phonon coupling. A qualitative sketch of the derivation of a general electron-
phonon Hamiltonian is provided here, while a more rigorous derivation can
be found in Ref. [14].

The most general electron-phonon coupling Hamiltonian can be written
as follows:

\[ H = H_{el} + H_{ph} + H_{el-ph} + H_{el-el}. \] (1.1)

Here the first two terms represent the the kinetic energy of the electrons
and phonons, respectively, and the last two terms give the electron-phonon
coupling and electron-electron interactions in the system. For the single
electron (polaron) problem of interest to us, the spin of the electron is ir-
relevant and we suppress its index. Moreover, there are no el-el interactions
possible in this case.

A general lattice of linear oscillators can be written as

\[ H_{ph} = \sum_{q,\nu} \Omega_{\nu}(q) b_{q,\nu}^{\dagger} b_{q,\nu}, \] (1.2)

where \( b_{q,\nu}^{\dagger} \) and \( b_{q,\nu} \) are the creation and annihilation operators, respectively,
for a phonon with momentum \( q \) and branch index \( \nu \). The dispersion for each
1.2. Model of the Electron-Phonon Interaction

branch of phonons is $\Omega_\nu(q)$. The lattice potential felt by a single electron can be written as

$$V(r) = \sum_j v(r - R_j), \quad (1.3)$$

where $R_j$ are the positions of the lattice ions. To first order, the deviation of these lattice ions from their equilibrium positions $R_j^{(0)}$ will be small, so we can Taylor expand around $R_j^{(0)}$ to linear order and obtain

$$v(r - R_j) \simeq v(r - R_j^{(0)}) + \frac{\partial v(r - R_j)}{\partial R} \bigg|_{R_j = R_j^{(0)}} (R_j^{(0)} - R_j). \quad (1.4)$$

The combination of the kinetic energy of the electron with the periodic potential given by the first term of Eq. (1.4) produces the usual non-interacting Bloch states, and we can write

$$\mathcal{H}_{el} = \sum_k \varepsilon_k c_k^\dagger c_k, \quad (1.5)$$

where $c_k^\dagger$ and $c_k$ are the creation and annihilation operators, respectively, for an electron with momentum $k$, and the electron dispersion is given by $\varepsilon_k = \varepsilon(k)$. The second term of Eq. (1.4) gives rise to the electron-phonon interaction. The displacement $x_j = R_j^{(0)} - R_j$ can be written in terms of the phonon operators as $x_j \sim (b_j^\dagger + b_j)$, which allows us to write the electron-phonon interaction for a single phonon mode as

$$\mathcal{H}_{el-ph} = \sum_{i,j} g_{ij} c_i^\dagger c_i (b_j^\dagger + b_j), \quad (1.6)$$

where $g_{ij}$ contains the details of the electron-phonon coupling. Fourier transforming to $k$-space and allowing the possibility of several phonon modes, we obtain a general form for the electron-phonon coupling term:

$$\mathcal{H}_{el-ph} = \frac{1}{\sqrt{N}} \sum_{k,q,\nu} g_\nu(q) c_{k-q}^\dagger c_k (b_{q,\nu}^\dagger + b_{-q,\nu}), \quad (1.7)$$

where $N$ is the number of sites in the $d$-dimensional lattice, and all sums over momenta are over the first Brillouin zone (BZ).

There is a second effect of el-ph interactions, which is not captured by Eq. (1.7). As the on-site energy of the electron is changed because of the cloud forming around it [the physics described by Eq. (1.7)], the orbital occupied
by the electron changes its wavefunction, and therefore the hopping of the electron to neighbouring sites is also modified. In the linear limit, this gives rise to general terms of the type:

$$H'_{\text{el-ph}} = \sum_{i,j,k} t_{i,j,k} c_i^\dagger c_j (b_k^\dagger + b_k) + h.c.,$$

where the electron changes its location as a result of the interactions with phonons. By Fourier transforming this term, one obtains a Hamiltonian of the form

$$H = \sum_k \epsilon_k c_k^\dagger c_k + \sum_q \frac{1}{\sqrt{N}} \sum_{k,q} g_{\nu}(k,q) c_k^\dagger c_k \left( b_q^\dagger b_q + b_{-q}^\dagger b_{-q} \right),$$

where the el-ph coupling now depends on both the electron momentum and the phonon momentum. Such phonon-modulated hoping is very important in organic materials, polyacetylene being a very famous example, and it can lead to very interesting consequences [4].

The details of the electron-phonon interaction are included in the matrix element $g(k,q)$. In this work we will consider cases where $g(k,q) = g(q)$, i.e. cases where the el-ph coupling is dependent on the phonon momentum only, and we will always consider coupling to a single phonon mode, although generalizations to many phonon modes are possible [15]. We can further simplify the model by recognizing that in ionic crystals the most dominant el-ph coupling is to optical phonons. Optical phonons correspond to the crystal vibrating out of phase, and in an ionic crystal this motion sets up a dipole polarization field along the axes of vibration because the lattice sites are either positively or negatively charged. The resulting electric field induced by the optical phonons is primarily responsible for the coupling of the electron to the lattice.

Therefore, by considering a single branch of Einstein optical phonons of constant frequency $\Omega$, the general electron-phonon coupling model that we consider in this thesis has the following form in momentum space:

$$H = \sum_k \epsilon_k c_k^\dagger c_k + \sum_q \frac{1}{\sqrt{N}} \sum_{k,q} g_q c_k^\dagger c_k \left( b_q^\dagger b_q + b_{-q}^\dagger b_{-q} \right).$$

Equation (1.10) reduces to the Holstein model [16, 17] when $g_q$ is simply a constant, but this general Hamiltonian also covers more complicated electron-phonon couplings such as those found in the Rashba-Pekkar [18, 19] and Fröhlich [20] models, where $g_q \sim 1/q$, and the breathing-mode Hamiltonian [21, 22], where $g_q \sim \sin(q/2)$ in 1D.
1.3 The Green’s Function

Before further discussing the electron-phonon coupling models studied in this work, we first review the Green’s function formalism that we apply to study such models. The quantity of interest to us is the Green’s function of the single dressed particle, or polaron. This one-particle Green’s function for an M-electron system is defined as [14]:

\[ G(k, \tau) = -i\langle \Phi_M | T[c^\dagger_k(\tau)c_k(0)] | \Phi_M \rangle, \quad (1.11) \]

where \( |\Phi_M \rangle \) is the M-electron ground state, \( T \) is the time ordering operator, and

\[ c_k(\tau) = e^{iH\tau}c_k e^{-iH\tau} \quad (1.12) \]

is the destruction operator for an electron of momentum \( k \) in the Heisenberg representation. We have also set \( \hbar = 1 \). To study the spectrum of a single dressed electron we set \( M = 0 \) and denote \( |\Phi_0 \rangle = |0\rangle \), where \( |0\rangle \) is the vacuum. Equation (1.11) then reduces to

\[ G(k, \tau) = -i\Theta(\tau)|c_0 e^{-iH\tau}c^\dagger_0|0\rangle, \quad (1.13) \]

where \( \Theta(\tau) \) is the Heaviside function, defined as

\[ \Theta(\tau) = \begin{cases} 1, & \tau > 0, \\ 0, & \tau < 0. \end{cases} \quad (1.14) \]

Therefore, we see that in the case of a single dressed electron only the retarded part contributes. This Green’s function gives the amplitude of probability that an electron introduced in the system at time zero and removed at a later time \( \tau \) leaves the system in its ground state.

The usefulness of this Green’s function becomes more apparent in the Lehmann representation [14]. By introducing the complete set of one-particle eigenstates and eigenenergies \( \{|\alpha\rangle\} \) and \( \{E_\alpha\} \), we can write the Green’s function as

\[ G(k, \omega) = \sum_\alpha \frac{|\langle \alpha | c^\dagger_k|0\rangle|^2}{\omega - E_\alpha + i\eta}. \quad (1.15) \]

Here we explicitly see that the poles of the Green’s function give the energy spectrum, while the residues give information on the nature of the eigenstates. Later on we will show that several other polaron properties can be calculated directly from this Green’s function, including the average number of phonons in the polaronic cloud and the effective mass of the electron.
1.4. The Holstein Model

In addition, the imaginary part of the Green’s function, or spectral weight, defined as

\[
A(k, \omega) = -\frac{1}{\pi} \text{Im} G(k, \omega) = \sum_{\alpha} |\langle \alpha | c^\dagger_k |0 \rangle|^2 \delta(\omega - E_{\alpha}),
\]

(1.16)

can be measured experimentally using angle-resolved (inverse) photoemission spectroscopy (ARPES/ARIPES) [23]. Strictly speaking, the spectral function in Eq. (1.16) would be measured using ARIPES by injecting an electron of momentum \( k \) into the system and analyzing the outgoing photon. However, ARPES in a full band will give the same spectrum with the energies “flipped”, i.e. \( A(k, -\omega) \).

1.4 The Holstein Model

Returning to the topic of electron-phonon coupling, we first describe the Holstein model [16, 17], the simplest Hamiltonian describing an electron on a lattice interacting with an optical phonon mode through on-site electron-phonon coupling. The Holstein model is given by the following Hamiltonian:

\[
\mathcal{H} = \sum_{k} \varepsilon_k c^\dagger_k c_k + \Omega \sum_{\mathbf{q}} b^\dagger_{\mathbf{q}} b_{\mathbf{q}} + \frac{g}{\sqrt{N}} \sum_{\mathbf{k},\mathbf{q}} c^\dagger_{\mathbf{k}-\mathbf{q}} c_{\mathbf{k}} \left( b^\dagger_{\mathbf{q}} + b_{\mathbf{-q}} \right).
\]

(1.17)

The units are such that \( \hbar = 1 \), and the notation is the same as the general Hamiltonian of Eq. (1.10), derived in Section 1.2. For the free electron dispersion we use nearest-neighbour hopping on a \( d \)-dimensional simple cubic lattice of constant \( a \) and a total of \( N \) sites, giving

\[
\varepsilon_k = -2t \sum_{i=1}^{d} \cos(k_i a).
\]

(1.18)

Here the electron-phonon coupling is a momentum independent on-site linear coupling

\[
\hat{V} = g \sum_{i} c^\dagger_i c_i (b^\dagger_i + b_i)
\]

(1.19)

between the electron and the phonons with a constant coupling constant \( g \), written in \( \mathbf{k} \)-space. All sums over momenta are over the BZ, \( k_i \in \left( \frac{-\pi}{a}, \frac{\pi}{a} \right) \), \( i = 1, \ldots, d \).

This model was originally proposed by Holstein to describe a molecular crystal with the vibrational degree of freedom perpendicular to the chain.
1.4. The Holstein Model

[16, 17]. Although the model appears quite simple, it can not be solved exactly. Throughout the years it has been widely studied using a variety of methods, both analytical and numerical, and it serves as an excellent starting point and a benchmark for numerical approaches and analytical techniques, such as the approximation presented in this work. Furthermore, it captures the key polaron physics, such as the crossover from a large polaron at weak couplings to a small polaron at large couplings.

1.4.1 Analytical Techniques

Of course, given the long history of this problem, many analytical techniques have been applied to study the Holstein model, with varying degrees of success. We first point out that the solution, and associated Green’s functions, are known exactly in two asymptotic limits. If there is no coupling, \( g = 0 \), then the Green’s function is that of the free electron:

\[
G_0(k, \omega) = \frac{1}{\omega - \varepsilon_k + i\eta}.
\]  

(1.20)

The ground state is at \( \varepsilon_0 = -2dt \) and the spectrum consists of a continuous band extending from \([-2dt, 2dt]\) (for the tight-binding model).

The so-called impurity limit, with \( t = 0 \), also has an exact solution, given by the Lang-Firsov formula [24]:

\[
G(\omega) = e^{-g^2/\Omega^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{g}{\Omega} \right)^{2n} \frac{1}{\omega + \frac{g^2}{\Omega} - n\Omega + i\eta}.
\]  

(1.21)

This can be viewed as the strong-coupling limit, since for \( t = 0 \), \( g \) becomes the important energy scale in the system. In this limit the electron is localized at one site in real space, therefore it is fully delocalized in k-space, and the Green’s function is independent of \( k \). The spectrum has the ground state (GS) at \( E_0 = -g^2/\Omega \) and an infinite sequence of equidistant levels spaced by \( \Omega \) above it (shown below). This is extremely different from the free-particle spectrum, and it is of considerable interest to understand not only how the ground state evolves from \(-2dt\) to \(-g^2/\Omega\) as the coupling \( g \) is increased, but also the evolution of all the higher-energy spectral weight from a continuous, finite-width band to an infinite set of discrete levels.

Naturally, one can extend the applicability of these results using perturbation theory. In the weak coupling regime one can use standard Rayleigh - Schrödinger (RS) perturbation theory. To first order this gives the lowest
1.4. The Holstein Model

energy for a state with total momentum $k$ as [25]:

$$E_{RS}^{H}(k) = \varepsilon_k - \frac{1}{N} \sum_q \frac{g^2}{\varepsilon_{k-q} + \Omega - \varepsilon_k}. \quad (1.22)$$

The summation is straightforward, and the result is

$$E_{RS}^{H}(k) = \varepsilon_k - g^2 \bar{g}_0 (\Omega - \varepsilon_k), \quad (1.23)$$

where

$$\bar{g}_0(\omega) = \frac{1}{N} \sum_k \frac{1}{\omega - \varepsilon_k + i\eta} \quad (1.24)$$

is the momentum average of the free propagator, given explicitly in Eqs. (2.47)-(2.49). While quite successful for small electron-phonon couplings, this result breaks down as the momentum $k$ of the electron is increased. This can be seen more clearly from the denominator of Eq. (1.22), which goes to zero as the energy of the electron approaches the energy of a free electron plus one phonon. For small couplings and larger momenta it is more appropriate to use the Self-Consistent Born approximation, discussed below.

At strong couplings one can again use perturbation theory. To demonstrate how this is done we will first explicitly derive the $t=0$ result, $E_0 = -g^2/\Omega$, already given above. In the limit of $t=0$ one can diagonalize the remaining two terms of Eq. (1.17) using the Lang-Firsov transformation [24, 26]. We transform the electron and phonon operators as

$$\tilde{c}_i = e^S c_i e^{-S}, \quad (1.25)$$

where

$$S = -\frac{g}{\Omega} \sum_i n_i \left( b_i^\dagger - b_i \right), \quad (1.26)$$

and $n_i = c_i^\dagger c_i$. Using the formula

$$\tilde{A} = e^S A e^{-S} = A + [S, A] + \frac{1}{2!} [S, [S, A]] + \cdots, \quad (1.27)$$

one obtains

$$\tilde{b}_i = b_i + \frac{g}{\Omega} n_i, \quad (1.28)$$

$$\tilde{c}_i = c_i \exp \left[ \frac{g}{\Omega} \left( b_i^\dagger - b_i \right) \right]. \quad (1.29)$$
Physically this corresponds to a shift of the equilibrium positions at ions where the electron is present. The transformed Hamiltonian can then be written as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_t$, where

$$\mathcal{H}_0 = -\frac{g^2}{\Omega} \sum_i \hat{n}_i^2 + \Omega \sum_i \hat{b}_i^\dagger \hat{b}_i$$

and

$$\mathcal{H}_t = -t \sum_{\langle i,j \rangle} \left( \hat{c}_i^\dagger \hat{c}_j \hat{X}_i^\dagger \hat{X}_j + h.c. \right).$$

Here $\hat{n}_i = \hat{c}_i^\dagger \hat{c}_i$, and we have also defined

$$X_i = e^{-\frac{g}{\Omega} (\hat{b}_i^\dagger - \hat{b}_i)}.$$ 

In the $t = 0$ limit $\mathcal{H}_t = 0$, and $E_0 = -g^2/\Omega$. This ground state is $N$-degenerate, corresponding to each of the $N$ sites of the lattice on which the electron can be localized. The $n^{th}$ excited state in this $t = 0$ limit corresponds to the creation of $n$ phonons in the system, giving an energy spectrum of the form seen in the Green’s function of Eq. (1.21).

The first and second order perturbation corrections can be calculated from $\mathcal{H}_t$. It turns out that the first order correction is exponentially suppressed, and that the second order correction is actually a more dominant term. These corrections can be calculated, and the full second order perturbation theory result is given by the following [22, 26]:

$$E_{\text{PT}}^H(\mathbf{k}) = -\frac{g^2}{\Omega} + \varepsilon_k e^{-g^2/\Omega^2} - \frac{\Omega t^2}{g^2} \left[ 1 + 2e^{-g^2/\Omega^2} \cos(2ka) \right].$$

The first order correction suggests the definition of an exponentially reduced nearest-neighbour hopping given by

$$t^* = te^{-g^2/\Omega^2},$$

which in turn gives rise to an exponentially increased polaron mass, given by

$$m^* = me^{g^2/\Omega^2}.$$ 

The number of phonons in the initial basis and the quasiparticle weight are also directly related to the ratio $g^2/\Omega^2$:

$$N_{\text{ph}} = \langle \hat{b}_i^\dagger \hat{b}_i \rangle = \langle (\hat{b}_i^\dagger - \frac{g}{\Omega})(\hat{b}_i - \frac{g}{\Omega}) \rangle = \frac{g^2}{\Omega^2}.$$
1.4. The Holstein Model

<table>
<thead>
<tr>
<th>order</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>1</td>
<td>2</td>
<td>10</td>
<td>74</td>
<td>706</td>
<td>8162</td>
<td>110410</td>
<td>1708394</td>
</tr>
<tr>
<td>SCBA</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>14</td>
<td>42</td>
<td>132</td>
<td>429</td>
</tr>
</tbody>
</table>

Figure 1.1: Comparison between the total number of diagrams of a given order in the proper self-energy $\Sigma(k, \omega)$ vs. the number of diagrams of a given order kept within SCBA.

$$Z_0 = |\langle X_i \rangle|^2 = e^{-g^2/\Omega^2}. \tag{1.37}$$

This is consistent with the idea of a small polaron at large electron-phonon couplings. In this regime the polaron gains an exponentially large mass while the lattice is strongly deformed over a small spatial range. This is in contrast to the large polaron found at small electron-phonon couplings. In this opposite extreme the electron is still very mobile and dressed by interactions extending over a large spatial range.

We briefly review here the other two simple (in terms of computational effort) analytical approximations for the Green’s function of the Holstein polaron available in the literature. The first is the self-consistent Born approximation (SCBA), which consists of summing exactly only the non-crossed diagrams. The percentage of diagrams kept decreases fast with increasing order, as shown in Fig. 1.1. If the coupling is small, the sum is dominated by the low order diagrams and SCBA works reasonably well. At strong coupling, the contribution of higher order diagrams becomes essential, and SCBA is expected to fail. In this approximation, the Green’s function is written in terms of a self-energy:

$$G_{SCBA}(k, \omega) = \frac{1}{\omega - \varepsilon_k - \Sigma_{SCBA}(\omega) + i\eta}, \tag{1.38}$$

with the self-consistency condition

$$\Sigma_{SCBA}(\omega) = g^2 N \sum_q G_{SCBA}(k - q, \omega - \Omega). \tag{1.39}$$

Note that $\Sigma_{SCBA}(\omega)$ is independent of $k$. This is a consequence of the simplicity of the Holstein model: if either the coupling $g$ or the dispersion $\Omega$ were functions of the phonon momentum $q$, the SCBA self-energy would depend explicitly on $k$ [21]. Because of this, the SCBA self-energy can be expressed as a function of the average of the free propagator over the BZ.
1.4. The Holstein Model

for the Holstein model and can be evaluated very efficiently. Rewriting the self-consistency condition above, one obtains the following:

\[
\Sigma_{\text{SCBA}}(\omega) = \frac{g^2}{N} \sum_q G_{\text{SCBA}}(k-q, \omega - \Omega) \]

\[
= \frac{g^2}{N} \sum_q \frac{1}{\omega - \varepsilon_{k-q} - \Omega - \Sigma_{\text{SCBA}}(\omega - \Omega) + i\eta} \] \hspace{1cm} (1.41)

\[
= \frac{g^2}{N} \sum_q G_0(k-q, \omega - \Omega - \Sigma_{\text{SCBA}}(\omega - \Omega)), \] \hspace{1cm} (1.42)

where we have made use of the fact that \(\Sigma_{\text{SCBA}}\) does not depend on momentum in the last step. Rewriting the self-energy in terms of \(\tilde{g}_0(\omega)\) [see Eq. (1.24)] one obtains

\[
\Sigma_{\text{SCBA}}(\omega) = g^2 \tilde{g}_0(\omega - \Omega - \Sigma_{\text{SCBA}}(\omega - \Omega)), \] \hspace{1cm} (1.43)

which can be expanded to:

\[
\Sigma_{\text{SCBA}}(\omega) = g^2 \tilde{g}_0(\omega - \Omega - g^2 \tilde{g}_0(\omega - 2\Omega - g^2 \tilde{g}_0(\omega - 3\Omega - \cdots))). \] \hspace{1cm} (1.44)

Another simple analytical approximation for the Green’s function of the Holstein model is the generalized Lang-Firsov (LF) expression \[27, 28\]. It is reminiscent of the Lang-Firsov expression of Eq. (1.21):

\[
G_{\text{LF}}(k, \omega) = e^{-g^2/\Omega^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{g}{\tilde{g}_0} \right)^{2n} \frac{1}{\omega - e^{-g^2/\Omega^2} \varepsilon_k + g^2/\Omega - n\Omega + i\eta}. \] \hspace{1cm} (1.45)

This expression is exact for both zero coupling \((g = 0)\) and zero bandwidth \((t = 0)\), however intermediary results are extremely poor, as we will demonstrate in Section 2.3.

We also note that it is customary to define a dimensionless coupling parameter \(\lambda\) as the ratio between the Holstein lattice deformation energy, \(E_0 = -g^2/\Omega\), and the electron half-bandwidth, \(W/2 = zt = -2dt\), where \(z\) is the coordination number and \(d\) is the dimension of the system:

\[
\lambda = \frac{E_0}{W/2} = \frac{g^2}{zt\Omega} = \frac{g^2}{2dt\Omega}. \] \hspace{1cm} (1.46)

We typically investigate couplings in the range \(0 < \lambda < 2\), and the intermediate couplings where the above analytical methods all fail occur when the lattice deformation energy and electron half-bandwidth are of the same order, \(i.e.\) when \(\lambda \sim 1\). Results obtained from the various analytical results described above will be presented in Section 2.3.
1.4.2 Numerical Methods

There is a large body of work devoted to solving the Holstein model using numerical methods, and we briefly review some of these methods here. The most widely used techniques are exact diagonalization methods (ED) [29–38] and various quantum Monte Carlo (QMC) methods [39–50]. ED is arguably the best controlled numerical tool for evaluating ground state and excited state properties for the Holstein model, although ED methods are limited because even for a finite lattice the Hilbert space is infinite because of the infinite number of possible phonon configurations, and some form of truncation is always required. In practice ED is usually run on small lattices due to memory restrictions, typically less than 20 sites. On the other hand, QMC methods can run on much larger lattices, on the order of 1000 sites or more, and for finite temperatures. Diagrammatic Monte Carlo [26, 51, 52] is of the most interest to us, as it calculates the Green’s functions in imaginary time by numerically summing all of the diagrams in the perturbational expansion. However, the analytical continuation from imaginary time limits the applicability of this method to ground state properties only. There have been several other numerical methods applied to the Holstein model as well. Variational methods (VM) [53–62], have been used in conjunction with ED methods, and combinations of ED with kernel polynomial [63, 64] and cluster perturbation methods [65, 66] have also been developed. The density matrix renormalization group (DMRG) [67, 68] and dynamic DMRG [69] are very accurate for systems on large lattices, but are limited to one dimension and low energy spectral properties, and dynamical mean-field theory (DMFT) can be applied for infinite-dimensional systems [70, 71].

Most of the numerical methods discussed above calculate only GS or low-energy properties, given the significant CPU time and numerical resources needed to calculate the whole spectrum. However, very recently, several sets of whole-spectrum results have become available [22, 48, 50, 62, 66], however only for a few points in the parameter space, and generally for low dimensions. A nice review of the applications of numerical methods to the Holstein model can also be found in Ref. [72].

1.5 The Breathing-mode Hamiltonian

In Chapter 4 we will study more complicated electron-phonon coupling models of the form given in Eq. (1.10). We will primarily use the 1D breathing-mode Hamiltonian as an example. The reason for studying the 1D breathing-mode Hamiltonian is two-fold. First, in its full 2D form, it describes lattice
1.5. The Breathing-mode Hamiltonian

vibrations in a CuO$_2$-like plane, where the motion of the O ions living on the
bonds connecting the Cu sites is the most important vibrational degree of
freedom [22], making it possibly relevant for the study of high-$T_c$ supercon-
ductors. The second reason is that exact diagonalization (ED) results [22]
have recently become available for its 1D analog, relevant for CuO chains.
In Chapter 4 these results serve as an excellent gauge of the accuracy of
the approximations presented in this work. Their availability is very for-
tunate because although there are many numerical results for the Holstein
model, it is only due to recent advancements in computational power that
more complicated electron-phonon coupling models, such as the breathing-
mode Hamiltonian, can be investigated numerically. We also mention that
this model has been studied using the self-consistent Born approximation
(SCBA) [21] but SCBA is known to be very poor for intermediate and large
coupling strengths, as illustrated in Section 2.3 for the Holstein model.

In the 1D breathing-mode model, one considers a chain with two inter-
laced sub-lattices, where the Cu sites which host the electron are indexed
by integer labels, and the O sites which host the phonons are indexed by
half-integer labels. The interaction term of the breathing-mode Hamiltonian
can be written in real space as

$$
\hat{V} = g \sum_i c_i^\dagger c_i \left( x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \right),
$$

(1.47)

where $x_{i\pm1/2} = b_{i\pm1/2}^\dagger + b_{i\pm1/2}$ describe the displacements of the O atoms
neighbouring the Copper atom at site $i$ that hosts the electron, and $g$ is a
constant describing the strength of the electron-phonon coupling ($g$ absorbs
the proportionality factors between the true displacement $x_{i\pm1/2}$ and the
phonon operator $b_{i\pm1/2}^\dagger + b_{i\pm1/2}$). Transforming into momentum space, the
1D breathing-mode Hamiltonian takes the form of Eq. (1.10) with

$$
g_q = -2ig\sin \frac{qa}{2}.
$$

(1.48)

In this model the electron motion is described by a tight-binding model
with the usual $\varepsilon_k = -2t \cos(ka)$, although our results can be applied for any
dispersion.

1.5.1 Analytical Techniques

As was the case with the Holstein model, one can use perturbation theory
to study the weak and strong coupling regimes of the breathing Hamilto-
nian. In the limit of zero coupling we again obtain the Green’s function of
The Breathing-mode Hamiltonian

the free electron and a ground state $\varepsilon_0 = -2dt$. To obtain the standard Rayleigh-Schrödinger (RS) first order perturbation theory result, valid for small electron-phonon couplings, one evaluates the familiar expression:

$$E_{\text{RS}}^{B}(k) = \varepsilon_k - \frac{1}{N} \sum_q \frac{g^2_q}{\varepsilon_{k-q} + \Omega - \varepsilon_k}. \quad (1.49)$$

For the 1D breathing-mode Hamiltonian one finds that

$$E_{\text{RS}}^{B}(k) = \varepsilon_k - 2g^2\bar{g}_0(\Omega - \varepsilon_k) - \frac{g^2}{t} \cos k \left[ 1 - (\Omega - \varepsilon_k)\bar{g}_0(\Omega - \varepsilon_k) \right], \quad (1.50)$$

where $\bar{g}_0(\omega)$ is defined in Eq. (1.24). This result has the same qualitative features as the result for the Holstein model, and while it is successful for smaller couplings, it breaks down for intermediate to large couplings and for large momenta $k$.

It is also instructive to calculate the strong-coupling perturbation theory result, as in the Holstein case. As before, one proceeds via the Lang-Firsov transformation given in Eq. (1.25), with

$$S = -\frac{g}{\Omega} \sum_i \bar{n}_i \left( x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \right). \quad (1.51)$$

The transformed operators are then:

$$\tilde{b}_{i+\frac{1}{2}} = b_{i+\frac{1}{2}} + \frac{g}{\Omega} \left( n_{i+1} - n_i \right) \quad (1.52)$$

and

$$\tilde{c}_i = c_i \exp \left[ \frac{g}{\Omega} \left( x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \right) \right]. \quad (1.53)$$

The Hamiltonian can again be written in the form $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_t$, where

$$\mathcal{H}_0 = -\frac{2g^2}{\Omega} \sum_i \bar{n}_i(\bar{n}_i - \bar{n}_{i+1}) + \Omega \sum_i \tilde{b}^\dagger_{i+\frac{1}{2}} \tilde{b}_{i+\frac{1}{2}} \quad (1.54)$$

and

$$\mathcal{H}_t = -t \sum_{\langle i,j \rangle} \left( \tilde{c}^\dagger_i \tilde{c}_j X_i^\dagger X_j + h.c. \right), \quad (1.55)$$

and we have defined

$$X_i = \exp \left[ -\frac{g}{\Omega} \left( x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} \right) \right]. \quad (1.56)$$
1.5. The Breathing-mode Hamiltonian

The physical meaning of this transformation is the same as in the Holstein case: the equilibrium positions of the the ions in the vicinity of the electron have been shifted due to the localized electron. The first and second order corrections to the ground state energy can again be obtained from $\mathcal{H}_t$, as in the Holstein case, and the final result in 1D is given by [22]:

$$E_{PT}^B (k) = -\frac{2g^2}{\Omega} + \varepsilon_k e^{-3g^2/\Omega^2} - \frac{\Omega t^2}{g^2} \left[ \frac{1}{3} + \frac{e^{-2g^2/\Omega^2}}{2} \cos(2ka) \right]. \quad (1.57)$$

This breathing-mode result has some key differences from the Holstein expression, given in Eq. (1.33). The first is that the lattice deformation energy for the 1D breathing-mode, $-\frac{2g^2}{\Omega}$, is twice that of the Holstein value. In fact, one can show that for a $d$-dimensional lattice this deformation energy is proportional to the number of neighbouring phonon sites, which is always one for the Holstein model, and equal to the coordination number $z = 2d$ for the breathing-mode. Therefore

$$E_0 = -\frac{zg^2}{\Omega} = -\frac{2dg^2}{\Omega} \quad (1.58)$$

for the $d$-dimensional breathing-mode Hamiltonian. As before, we define the dimensionless coupling constant as the ratio of the lattice deformation energy and half-bandwidth $W = -zt/2$. For the breathing-mode Hamiltonian this is the same in any dimension, and twice the Holstein value in 1D:

$$\lambda_B = \frac{zg^2}{\Omega} \frac{1}{zt} = \frac{g^2}{\Omega t}. \quad (1.59)$$

It is also illustrative to write Eqs. (1.33) and (1.57) in terms of their respective coupling constants, $\lambda$ and $\lambda_B$. Rewriting these equations one can show that

$$E_H(k) = -2\lambda t - 2te^{-\alpha\lambda} \cos(ka) - \frac{t}{\lambda} \left[ \frac{1}{3} + e^{-\alpha\lambda} \cos(2ka) \right] \quad (1.60)$$

and

$$E_B(k) = -2\lambda_B t - 2te^{-(3/2)\alpha\lambda_B} \cos(ka) - \frac{t}{\lambda_B} \left[ \frac{1}{3} + \frac{e^{-\alpha\lambda_B}}{2} \cos(2ka) \right], \quad (1.61)$$

where we have also defined the adiabatic ratio $\alpha = zt/\Omega$. Since the first order correction is exponentially suppressed, a comparison of the second order corrections for the same effective couplings ($\lambda = \lambda_B$) reveals that
1.6. The Highly Inhomogeneous Coupling Model

The highly inhomogeneous coupling (HIC) model, first proposed in Ref. [73], describes the coupling of an electron to a single phonon of momentum \( \mathbf{Q} \), through an el-ph coupling given by \( g_{\mathbf{Q}} = g_{\mathbf{Q}} \). In the notation of Eq. (1.10) this Hamiltonian takes the following form:

\[
\mathcal{H} = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + \Omega_{\mathbf{Q}} b_{\mathbf{Q}}^\dagger b_{\mathbf{Q}} + \frac{g}{\sqrt{N}} \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \left( b_{\mathbf{Q}}^\dagger + b_{\mathbf{Q}} \right). \tag{1.62}
\]

While this model may seem rather unphysical because it implies an infinitely-ranged coupling in real-space, there are examples of systems where \( g_{\mathbf{Q}} \) is peaked in the “corner” of the BZ. A good example is the breathing-mode Hamiltonian, described in the last section. Furthermore, having an exact solution to a model with such a long-range coupling allows one to draw conclusions about the behaviour of polarons in more realistic models possessing long-range coupling that may be difficult to solve, such as the Fröhlich model.

Although we will not study the physical properties of this model in detail in this work, the HIC model will provide us with some insight about the success of our approximations in models with long-range coupling. By gauging the accuracy of our approximations against the Holstein, breathing-mode, and HIC models, which possess local, short-range, and infinite-range coupling, respectively, we aim to more fully understand the applicability of
1.7 Experimental Evidence for Polarons

The importance of polaronic physics has been well established and accepted in materials such as fullerenes, polymers, nanotubes and manganites [2–5]. There are also materials, such as Beryllium, that have surface states possessing significantly stronger polaronic effects at the surface than in the bulk [2, 74–76]. We will study the effect of a surface on a Holstein polaron in Chapter 5. However, in the cuprates the relevance of polaronic effects is still open to debate. The understanding of the microscopic pairing mechanism in

Figure 1.2: The kink in the quasiparticle dispersion, as measured by Lanzara et al. in Ref. [8]. Panels (a)-(e) show the quasiparticle dispersion for various high-$T_c$ compounds at different doping levels $\delta$ and temperature. The solid arrows show the relevant phonon energy, located in the vicinity of the kink in the dispersion, and the dashed black lines are a guide to the eye. Reprinted by permission from Macmillan Publishers Ltd: Nature [8], copyright 2001.

our approximation. More details with regard to the HIC model are given in Section 4.2.4.
1.7. Experimental Evidence for Polarons

High-$T_c$ superconductivity has remained an unsolved problem since its discovery over twenty years ago. While phonon exchange drives the formation of Cooper pairs in conventional superconductivity, there is no consensus on the origin of the pairing mechanism in the high-$T_c$ superconductivity. In the last several years there has been intense debate over the origin of a kink found in the quasiparticle dispersion (see Fig. 1.2), measured by ARPES to be in the range of 50-80 meV. While it is accepted that this kink is a signature of an electron coupling to a bosonic mode, there is still debate on whether this bosonic mode is phononic or magnetic in origin [77].

Evidence in favour of polarons in the cuprates has been given, for e.g., by Lanzara et al. [8] and Shen et al. [9]. We present a figure from Lanzara et al. in Fig. 1.2, which illustrates the kink in the quasiparticle dispersion for various high-$T_c$ compounds at different doping levels and temperatures. Lanzara et al. argue that this quasiparticle kink can be explained by a coupling to phonons with an energy in the range of 50-80 meV, as indicated by the large solid arrows in the figure.

In Fig. 1.3 we show data from Shen et al., taken from Ref. [9]. In their work they also argue for the presence of polarons in high-$T_c$ superconductors. They claim that the relevant el-ph coupling in cuprates is in the strong-coupling regime, consistent with the Franck-Condon (Lang-Firsov) spectrum shown in Fig. 1.3(a), and that the reason that they do not see a quasiparticle peak in the cuprates is because they are in the strong polaronic limit where the quasiparticle weight of the ground state is extremely small. They claim that this ground state energy is 500 meV below the main spectral features shown in Fig. 1.3(b). Taking the approximate parameters from their paper, $t = 100$ meV and $\Omega = 50$ meV, we vary the el-ph coupling $g$ in an attempt to find the best fit to their data, using the results of our momentum average (MA) approximation that will be presented in detail later in this work. In Fig. 1.4 we show that with $g = 250$ meV we can obtain a ground state having a separation of approximately 500 meV from the main spectral features. This choice for $g$ corresponds to a dimensionless coupling constant of $\lambda = 6.25$ [see Eq. (1.46)] and is in the strong el-ph coupling regime. The significance of the result in Fig. 1.4 is that the broad peak of Fig. 1.3 is consistent with the Gaussian envelope fit to the MA data in Fig. 1.4, and that this Gaussian envelope has the correct energy separation of 450 meV between the peak of the spectral weight and onset of the spectral weight, denoted by $A$ and $B$ respectively in Figs. 1.3(b) and 1.4. The ground state found within MA is marked with a blue arrow and is only visible on a logarithmic scale, shown in the inset of Fig. 1.4. This ground state has an extremely small weight and it would certainly be lost in the
1.7. Experimental Evidence for Polarons

![Figure 1.3: Further evidence for the presence of polarons in high-$T_c$ superconductors, as shown by Shen et al. in Ref. [9] a) The rightmost picture illustrates the Franck-Condon (Lang-Firsov) broadening model. b) Measured spectrum of $\text{Ca}_2\text{CuO}_2\text{Cl}_2$, fit to a Lorentzian spectral function (dashed) and Gaussian (red or gray). Comparison with $\text{Sr}_2\text{RuO}_4$ is shown (thin black). The inset of panel (b) shows the spectrum of polarons in a $\text{H}_2$ molecule. Reprinted figure with permission from Ref. [9]: Shen et al., Phys. Rev. Lett. 93, 267002 (2004). Copyright 2004 by the American Physical Society.](image)

noise background of ARPES measurements, in agreement with the claims of Shen et al. in Ref. [9]. However, the extremely large el-ph coupling required to fit our data to that of Shen et al. is unbelievably high and we do not assert that these single electron results obtained from the Holstein model are an adequate description of the true physical system. Furthermore, we have not addressed the temperature dependence or asymmetry of the measured spectrum. Nevertheless, it is interesting that a similar spectrum to that measured by Shen et al. can be obtained using the MA approximation and a simple model for the el-ph interaction.
1.8 Motivation

In this Introduction we have reviewed the important role of electron-phonon coupling in physical systems and qualitatively derived a general Hamiltonian that describes the coupling of an electron to its lattice. Even the simplest of such el-ph models, the Holstein model, cannot be solved exactly and has been the subject of significant analytical and numerical work over the last several decades. More realistic models such as the breathing-mode Hamiltonian are even more difficult to solve, as perturbation theory fails to describe the interesting physics that occurs at intermediate couplings, and numerical techniques are often limited to low-dimensionality and/or ground state properties only, not to mention requiring even more intensive computational resources than the Holstein model. Therefore, it is of obvious interest to find an analytical approximation for the Green’s function that is simple to estimate, so that the whole parameter space can be studied easily, but also with high accuracy. This is precisely the topic of this work. The approximation developed in this work is called the momentum average (MA) approxima-
1.8. Motivation

tion; its essence consists of analytically summing all the diagrams in the diagrammatic expansion, but with each diagram simplified in a certain way such that the full summation can be performed. It does not suffer from the shortcomings of the analytical techniques discussed in Sections 1.4.1 and 1.5.1, as it reproduces both the zero coupling and zero bandwidth limits exactly, and it is also highly accurate for intermediate couplings, unlike any of the other analytical approaches. On the other hand, it does not require highly sophisticated numerical algorithms and significant numerical effort like the existing numerical techniques. In the later chapters, we show how MA can be generalized to model Hamiltonians equal to and beyond the scope of state-of-the-art numerical techniques, further demonstrating the utility of this approximation. We also show how MA can be systematically improved. Although the accuracy at the basic level of approximation is very good, it is found that some key physical features are not reproduced by the zeroth level MA approximation. By developing improvements to MA we improve its accuracy everywhere and remedy these shortcomings. This also allows us to better understand the meaning of the MA approximation and apply it to more general Hamiltonians.

This work is organized as follows. In Chapter 2 we introduce the MA approximation and use it to study the Holstein model. A detailed analysis of its accuracy is presented, through diagrammatics and spectral weight sum rules, and its accuracy is further gauged through comparisons to existing numerical data. We also discuss the shortcomings of the zeroth-level MA approximation, and then in Chapter 3 we show how the approximation can be systematically improved, effectively remedying these problems. In Chapter 4 we apply the MA approximation to the more general class of el-ph Hamiltonians described by Eq. (1.10), presenting both a straightforward extension of the previous approximation, and also an extremely accurate variational approach, using the 1D breathing-mode Hamiltonian as an explicit example. In Chapter 5 we return to the Holstein model, but here we look at the effect of a surface on the spectra of Holstein polarons. In this case of broken translational invariance we find significant differences from the bulk properties, demonstrating that the interpretation of surface measurements is not straightforward. Finally, Chapter 6 contains our summary and conclusions.
Chapter 2

The Holstein Polaron

In this chapter we develop the momentum average (MA) approximation, a new analytical approximation that will serve as the tool for investigating the electron-phonon coupling models studied in this work. We will first apply the MA approximation to the Holstein model, extensively discussing the meaning and accuracy of the approximation, and we will then compare the results produced by MA to available numerical data to further gauge its accuracy. A detailed account on the origin of the Holstein model and a summary of analytical and numerical approaches that have been previously applied can be found in Section 1.4.

Recall that the Holstein model is defined as [see Eq. (1.17)]:

\[ H = \sum_{k} \varepsilon_{k} c_{k}^\dagger c_{k} + \Omega \sum_{q} b_{q}^\dagger b_{q} + \frac{g}{\sqrt{N}} \sum_{k,q} c_{k}^\dagger c_{k}^\dagger \left( b_{q}^\dagger + b_{-q} \right). \]  

(2.1)

The first term in the kinetic energy of the electron, where \( c_{k}^\dagger \) and \( c_{k} \) are the electron creation and annihilation operators, respectively, and \( \varepsilon_{k} \) is the dispersion for nearest-neighbour hopping on a \( d \)-dimensional simple cubic lattice of constant \( a \) and a total of \( N \) sites, given by Eq. (1.18). The second term describes a branch of Einstein optical phonons with frequency \( \Omega \), where the phonon creation and annihilation operators are given by \( b_{q}^\dagger \) and \( b_{q} \), respectively. The third term represents a momentum independent on-site linear coupling \( g \sum_{i} c_{i}^\dagger c_{i} \left( b_{i}^\dagger + b_{i} \right) \) between the electron and the phonons with a constant coupling constant \( g \), written in \( k \)-space. All sums over momenta are over the first BZ, \( k_{i} \in (\frac{-\pi}{a}, \frac{\pi}{a}] \), \( i = 1, \ldots, d \).

As discussed in Section 1.8, our goal is to obtain a highly accurate and easy to use approximation that is accurate over all of parameter space. The application of the MA approximation to the Holstein model presented here will serve as a basis for the future chapters, where we systematically improve the approximation and increase its range of applicability to more general el-phon models.

This chapter is organized as follows. In Section 2.1, we derive the equations of motion for the Holstein polaron and introduce the momentum average approximation. Its physical meaning and accuracy is discussed in
2.1. The Momentum Average Approximation

Section 2.2, where we also calculate its corresponding spectral weight sum rules. We show that MA satisfies exactly the first 6 sum rules, but more importantly, it remains highly accurate for higher order sum rules. This is a strong argument in favor of its accuracy. The accuracy is gauged in more detail in Section 2.3, where we compare the MA predictions against the SCBA and generalized LF predictions, and also against a host of numerical results. This further demonstrates that the MA approximation is remarkably accurate for all parameter values, especially given its simplicity. In Section 2.3 we also present several new results for various other properties of the Holstein polaron. Finally, Section 2.4 contains our summary and conclusions for the application of the MA approximation to the Holstein model.

\section{The Momentum Average Approximation}

\subsection{The Exact Solution}

Although exact solutions for el-ph problems are generally not obtainable in closed form, one can write down their formal solutions in terms of an infinite set of coupled equations involving related (higher-order) Green’s functions. We use the equation of motion technique to generate this infinite hierarchy of coupled equations for an infinite number of related Green’s functions, and derive it here for the Holstein polaron.

In the frequency domain, this approach is equivalent to using repeatedly Dyson’s identity \( \hat{G}(\omega) = \hat{G}_0(\omega) + \hat{G}(\omega) \hat{V} \hat{G}_0(\omega) \), which holds for any Green’s operators \( \hat{G}(\omega) = [\omega - \hat{\mathcal{H}} + i\eta]^{-1}, \hat{G}_0(\omega) = [\omega - \hat{\mathcal{H}}_0 + i\eta]^{-1} \) and for any Hamiltonian written in the form \( \hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{V} \). As is customary, we take \( \hat{V} \) to be the el-ph interaction in Eq. (2.1). Applying Dyson’s identity once, we obtain:

\begin{equation}
G(k, \omega) = G_0(k, \omega) \left[ 1 + \frac{g}{\sqrt{N}} \sum_{q_1} F_1(k, q_1, \omega) \right],
\end{equation}

where

\begin{equation}
G_0(k, \omega) = \langle 0 | c_k \hat{G}_0(\omega) c_k^\dagger | 0 \rangle = \frac{1}{\omega - \varepsilon_k + i\eta}
\end{equation}

is the free particle Green’s function. We made use of the equality \( \hat{V} c_k^\dagger | 0 \rangle = \frac{g}{\sqrt{N}} \sum_q c_{k-q}^\dagger b_q^\dagger | 0 \rangle \) and defined a new Green’s function:

\begin{equation}
F_1(k, q_1, \omega) = \langle 0 | c_k \hat{G}(\omega) c_{k-q_1}^\dagger \hat{b}_q^\dagger | 0 \rangle.
\end{equation}
2.1. The Momentum Average Approximation

The Green’s function $F_1(k, q_1, \omega)$ is related to the amplitude of probability to start with the electron and a phonon at the initial time, and find only the electron in the system at the final time. Its own equation of motion relates back to $G(k, \omega)$ but also to a new Green’s function with two phonons initially. In general, if we define:

$$F_n(k, q_1, \ldots, q_n, \omega) = \langle 0 | c_k^\dagger G(\omega) c_{k - q_T} b_{q_1}^\dagger \ldots b_{q_n}^\dagger | 0 \rangle,$$

where $q_T = \sum_{i=1}^n q_i$ is the total momentum of the $n$ initial phonons, using Dyson’s identity we find its equation of motion to be ($n \geq 1$):

$$F_n(k, q_1, \ldots, q_n, \omega) = \frac{g}{\sqrt{N}} G_0(k - q_T, \omega - n\Omega)$$

$$\times \left[ \sum_{i=1}^{n} F_{n-1}(k, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, \omega) + \sum_{q_{n+1}} F_{n+1}(k, q_1, \ldots, q_{n+1}, \omega) \right],$$

i.e. related to the Green’s functions with $n - 1$ and $n + 1$ initial phonons. Eqs. (2.2) and (2.6) form the exact infinite hierarchy of coupled equations whose solution is the Holstein polaron Green’s function $G(k, \omega) = F_0(k, \omega)$.

Observing from Eqs. (2.6) that all of these generalized Green’s functions $F_1, F_2, \ldots$ must be proportional to $G(k, \omega)$, we can recast our equations into a more convenient form by defining

$$f_n(q_1, \ldots, q_n) = \frac{N^{n/2} g^n F_n(q_1, \ldots, q_n)}{G(k, \omega)},$$

where we have also introduced the shorthand notation $f_n(k, q_1, \ldots, q_n, \omega) \equiv f_n(q_1, \ldots, q_n)$ (i.e., the $k$ and $\omega$ dependence of these functions is implicitly assumed from now on). In this notation, Eqs. (2.2) and (2.6) become

$$G(k, \omega) = G_0(k, \omega) \left[ 1 + \frac{1}{N} \sum_{q_1} f_1(q_1) G(k, \omega) \right]$$

and

$$f_n(q_1, \ldots, q_n) = G_0(k - q_T, \omega - n\Omega)$$

$$\times \left[ g^2 \sum_{i=1}^{n} f_{n-1}(\ldots, q_{i-1}, q_{i+1}, \ldots) + \frac{1}{N} \sum_{q_{n+1}} f_{n+1}(q_1, \ldots, q_{n+1}) \right],$$
2.1. The Momentum Average Approximation

\[ \Sigma(k, \omega) = \sum \left( \frac{1}{N} \sum_{q_1} f_1(q_1) \right) + \ldots \]

Figure 2.1: The diagrammatical expansion for the self-energy \( \Sigma(k, \omega) \).

respectively. The solution can be written in the standard form

\[ G(k, \omega) = \frac{1}{\omega - \varepsilon_k - \Sigma(k, \omega) + i\eta}, \tag{2.10} \]

where the self-energy is given by

\[ \Sigma(k, \omega) = \frac{1}{N} \sum_{q_1} f_1(q_1). \tag{2.11} \]

This self-energy can also be expanded in terms of an infinite set of diagrams, as shown in Fig. 2.1.

Of course this system can be solved trivially in the limit of \( g = 0 \), in which case \( G(k, \omega) = G_0(k, \omega) \) directly from Eq. (2.2). An exact solution equivalent to the Lang-Firsov limit must also exist if \( t = 0 \). Indeed, in this limit all Green’s functions become independent of all momenta, and Eqs. (2.8) and (2.9) can be written as:

\[ G(\omega) = G_0(\omega) \left[ 1 + f_1(\omega)G(\omega) \right] \tag{2.12} \]

and

\[ f_n(\omega) = G_0(\omega - n\Omega) \left[ ng^2 f_{n-1}(\omega) + f_{n+1}(\omega) \right], \tag{2.13} \]

respectively. To simplify the following steps we suppress functional notation and write

\[ f_n = \alpha_n f_{n-1} + \beta_n f_{n+1}, \tag{2.14} \]

where

\[ \alpha_n = ng^2 G_0(\omega - n\Omega), \quad \beta_n = G_0(\omega - n\Omega). \tag{2.15} \]

On physical grounds we expect that \( f_{n+1} \) goes to zero for sufficiently large \( n \) since these higher order processes become less and less likely as \( n \to \infty \). We can then write \( f_n = \alpha_n f_{n-1} \) at a sufficiently large cutoff \( n \). Knowing \( f_n \) in terms of \( f_{n-1} \), we can write \( f_{n-1} \) in terms of \( f_{n-2} \) and so on. After \( n - 1 \)
such iterations we end up with the following expression:

\[ f_1 = \frac{\alpha_1}{\alpha_2 \beta_1} f_0. \]  

(2.16)

Recalling that \( f_0 = G = G_0(1 + f_1 G) \), one can easily solve for \( G \). Restoring the original notation we obtain the following continued fraction form for the Green’s function \( G(\omega) \):

\[ G(\omega) = \frac{G_0(\omega)}{1 - \frac{g^2 G_0(\omega) G_0(\omega - \Omega)}{1 - \frac{2g^2 G_0(\omega - \Omega) G_0(\omega - 2\Omega)}{1 - \cdots}}}, \]

(2.17)

where \( G_0(\omega) = (\omega + i\eta)^{-1} \). Relaxing the condition that \( f_{n+1} \to 0 \) and taking the continued fraction to infinite order the above expression becomes exact and equal to the Lang-Firsov result of Eq. (1.21).

The exact hierarchy of coupled equations given by Eqs. (2.8) and (2.9) can in principle be solved in the general case of finite \( g \) and finite \( t \) by iteratively solving for \( f_1 \), then \( f_2 \), etc., and removing them from this coupled system. It is straightforward to verify that the solution obtained in this case is the diagrammatic expansion, which can be partially re-summed to the expected form:

\[ G(k,\omega) = \frac{1}{[G_0(k,\omega)]^{-1} - \Sigma(k,\omega)}, \]

(2.18)

where the self-energy \( \Sigma(k,\omega) \) is the sum of all proper self-energy diagrams, the first few of which are shown in Fig. 2.1. While this solution as a sum of an infinite number of diagrams is exact, it is clearly not very useful if the sum cannot be performed, as is the case here. One typical strategy in such cases is to sum only a subset of these diagrams (e.g. the non-crossed ones, in the Self-Consistent Born Approximation). This is reasonable when one can argue that the diagrams kept contribute much more than the neglected diagrams, which is not the case for SCBA in this problem. We propose a new strategy, explained below, that allows us to find an approximative solution of these equations in the case of finite \( t \) and finite \( g \).

### 2.1.2 The Approximation

To obtain an approximate solution for the case of finite hopping and finite coupling we proceed as follows. We first note that \( \Sigma(k,\omega) \) depends on \( f_1(q_1) \)
2.1. The Momentum Average Approximation

only through its average over the BZ [see Eq. (2.11)]. This suggests that we define momentum averaged Green’s functions given by

$$\mathcal{F}_n = \frac{1}{N^n} \sum_{\mathbf{q}_1, \ldots, \mathbf{q}_n} f_n(\mathbf{q}_1, \ldots, \mathbf{q}_n),$$  \hspace{1cm} (2.19)

where all sums are over the BZ, and we have again suppressed the $\mathbf{k}$ and $\omega$ labels. Then Eq. (2.8) can be written in terms of $\mathcal{F}_1$ as

$$G(\mathbf{k}, \omega) = G_0(\mathbf{k}, \omega) \left[ 1 + \mathcal{F}_1 G(\mathbf{k}, \omega) \right]. \hspace{1cm} (2.20)$$

We would also like to write the recurrence relations for $f_n(\mathbf{q}_1, \ldots, \mathbf{q}_n)$ in terms of $\mathcal{F}_n$. The first term on the right hand side of Eq. (2.9) can be written in terms of $\mathcal{F}_{n-1}$ exactly, however the second term requires an approximation in order to write it in terms of $\mathcal{F}_{n+1}$. We choose this approximation to be

$$\frac{1}{N^{n+1}} \sum_{\mathbf{q}_1, \ldots, \mathbf{q}_{n+1}} G_0(\mathbf{k} - \mathbf{q}_T, \omega - n\Omega) f_{n+1}(\mathbf{q}_1, \ldots, \mathbf{q}_{n+1}) \approx \bar{g}_0(\omega - n\Omega) \mathcal{F}_{n+1}, \hspace{1cm} (2.21)$$

where

$$\bar{g}_0(\omega) = \frac{1}{N} \sum_{\mathbf{k}} G_0(\mathbf{k}, \omega) \hspace{1cm} (2.22)$$

is the free propagator momentum-averaged over the BZ. One justification for replacing $G_0(\mathbf{k} - \mathbf{q}_T, \omega - n\Omega)$ by its momentum average $\bar{g}_0(\omega - n\Omega)$ in Eq. (2.21) is that $\mathbf{q}_T = \sum_{i=1}^n \mathbf{q}_i$ takes, with equal probability, any value in the BZ. Moreover, in the impurity limit $t = 0$ where all Green’s functions are momentum independent, Eq. (2.21) is exact. This suggests that our approach should be reasonable at least in the strong-coupling limit $t \ll g$. In a more practical sense, this approximation allows us to write $\mathcal{F}_n$ in terms of $\mathcal{F}_{n-1}$ and $\mathcal{F}_{n+1}$ only, which is what we require to be able to obtain an analytical expression for $G(\mathbf{k}, \omega)$. We discuss the meaning and consequences of this approximation in more detail below.

After the approximation of Eq. (2.21), Eqs. (2.9) become:

$$\mathcal{F}_n = \bar{g}_0(\omega - n\Omega) \left[ n^2 \mathcal{F}_{n-1} + \mathcal{F}_{n+1} \right]. \hspace{1cm} (2.23)$$

Together with Eq. (2.20) these simplified recurrence relations can be solved as in the $t = 0$ case. We find:

$$G(\mathbf{k}, \omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} - \Sigma_{\mathcal{MA}^{(0)}}(\omega) + i\eta}, \hspace{1cm} (2.24)$$

where $\Sigma_{\mathcal{MA}^{(0)}}(\omega) = \frac{1}{N} \sum_{\mathbf{k}} \Sigma(\mathbf{k}, \omega)$ is the momentum-averaged self-energy.
2.2. The Meaning of the MA Approximation

where the self-energy is, within this approximation,

$$\Sigma_{\text{MA}(0)}(\omega) = g^2 A_1(\omega), \quad (2.25)$$

and we have defined the following infinite continued fractions:

$$A_n(\omega) = \frac{n\bar{g}_0(\omega - n\Omega)}{1 - g^2\bar{g}_0(\omega - n\Omega)A_{n+1}(\omega)}$$

$$= \frac{n\bar{g}_0(\omega - n\Omega)}{1 - \frac{(n + 1)g^2\bar{g}_0(\omega - n\Omega)\bar{g}_0(\omega - (n + 1)\Omega)}{1 - \cdots}}, \quad (2.26)$$

This is the main result of this chapter. This is the lowest level of the MA approximation, which we call MA\(^{(0)}\). We will derive the systematic improvements to MA, denoted by MA\(^{(n)}\) with \(n = 0, 1, 2, \ldots\), in Chapter 3. Any reference to the MA approximation without a superscript is meant to refer to the MA\(^{(0)}\) level of approximation.

As discussed, the MA approximation becomes exact in the limit of zero hopping \((t = 0)\), where \(\bar{g}_0(\omega) \rightarrow \bar{G}_0(\omega) = (\omega + i\eta)^{-1}\), but also for zero coupling, \(g = 0\). In the following sections we show that the range of validity of this approximation extends well beyond these asymptotic limits, and that in fact the MA expression is reasonably accurate over the entire parameter space. We also note that although the MA self-energy looks similar to the DMFT self-energy \([70, 71]\) this is in fact a very different approximation. This issue is discussed in Appendix A.

Finally, a comparison of the MA result to the SCBA expression of Eq. (1.44) illustrates that both MA and SCBA require similar computational effort. However, in the following sections we will illustrate that MA is much more accurate over all of parameter space.

2.2 The Meaning of the MA Approximation

2.2.1 Diagrammatics

To understand the diagrammatical meaning of the MA approximation we expand Eq. (2.25) in powers of \(g^2\):

$$\Sigma_{\text{MA}(0)}(\omega) = g^2\bar{g}_0(\omega - \Omega) + g^4 \left[2\bar{g}_0^2(\omega - \Omega)\bar{g}_0(\omega - 2\Omega)\right]$$

$$+ g^6 \left[4\bar{g}_0^3(\omega - \Omega)\bar{g}_0^2(\omega - 2\Omega) + 6\bar{g}_0^2(\omega - \Omega)\bar{g}_0(\omega - 2\Omega)^2 + 6\bar{g}_0^2(\omega - \Omega)\bar{g}_0(\omega - 3\Omega)\right] + \mathcal{O}(g^8) \quad (2.27)$$
2.2. The Meaning of the MA Approximation

and analyze the various terms. First, one can verify that the MA approx-
imation generates the correct number of proper self-energy diagrams to all
orders. Indeed, there is one term of order $g^2$, two terms of order $g^4$, 4+6=10
terms of order $g^6$, and so on (see Fig. 1.1). Moreover, to each of these
terms we can associate an MA diagram. These have the same topology
as the exact proper self-energy diagrams. The difference is that each free
propagator $G_0(k,\omega)$ in the exact self-energy diagrams is replaced with a
momentum averaged free propagator $\bar{g}_0(\omega)$ (with the correct frequency) in
each MA diagram.

Using Eq. (2.22), the first-order self-energy diagram is (see Fig. 2.1):

$$\Sigma^{(1)}(k,\omega) = \frac{g^2}{N} \sum_{q} G_0(k - q, \omega - \Omega) = g^2 \bar{g}_0(\omega - \Omega), \quad (2.28)$$

i.e. $\Sigma^{(1)}(k,\omega) = \Sigma^{(1)}_{\text{MA}/0}(\omega)$, and thus MA is exact to first order. Differences
appear from the second order diagrams, where the two exact contributions
(see Fig. 2.1):

$$\frac{g^4}{N^2} \sum_{q_1, q_2} G_0(k - q_1, \omega - \Omega)G_0(k - q_1 - q_2, \omega - 2\Omega) \times [G_0(k - q_1, \omega - \Omega) + G_0(k - q_2, \omega - \Omega)] \quad (2.29)$$

are replaced, within MA, by two equal contributions:

$$2g^4 \bar{g}_0^2(\omega - \Omega)g_0(\omega - 2\Omega) = \frac{2g^4}{N^3} \sum_{q_1, q_2, q_3} G_0(q_1, \omega - \Omega) \times G_0(q_2, \omega - 2\Omega)G_0(q_3, \omega - \Omega). \quad (2.30)$$

Comparing Eq. (2.29) to Eq. (2.30), we see that the MA self-energy di-
agrams have the correct number of free propagators with the correct fre-
quencies, however the momenta of the free propagators are uncorrelated
and individually averaged over. It is as if there is no connection between
the momentum carried by a cloud phonon when it is emitted and when it
is reabsorbed by the electron. Precisely the same holds for all higher order
self-energy diagrams.

To gain a better understanding of the difference between the exact dia-
grams and the MA diagrams, let us further analyze the dependence on $t$ of
Eqs. (2.29) and (2.30). For $t = 0$ the expressions are identical, because the
free propagators become independent of momenta and, as already discussed, MA becomes exact. For finite $t$, we expand each free propagator as:

$$G_0(k, \omega) = G_0(\omega) \left[ 1 + \varepsilon_k G_0(\omega) + (\varepsilon_k G_0(\omega))^2 + \cdots \right], \quad (2.31)$$

where $G_0(\omega) = (\omega + i\eta)^{-1}$. Inserting this expansion into Eqs. (2.29) and (2.30) and collecting powers of $t$ yields the following. All $O(t)$ terms vanish in both the exact and the MA diagrams because they are proportional to a $\sum_q \varepsilon_q = 0$. In fact, all odd-order powers in $t$ vanish because $\sum_q \varepsilon_q^{2n+1} = 0$. Next, consider terms of order $t^2$. Such terms arise either from expanding one free propagator to $O(t^2)$, or from expanding two different free propagators to $O(t)$. The former case leads to the same result for both the exact and the MA diagrams, and we obtain 6 contributions proportional to $(1/N) \sum_q \varepsilon_q = 2 dt^2$. The latter case, however, reveals a difference. Five of the six $O(t^2)$ such contributions from the exact diagrams vanish because they involve propagators carrying different momenta, and, for example,

$$\sum_{q_1, q_2} \varepsilon_{k-q_1} \varepsilon_{k-q_1-q_2} = 0.$$ 

The exception comes from the two outside free propagators of the non-crossed diagram, which carry the same momentum and result into another $(1/N) \sum_{q_1} \varepsilon_{k-q_1}^2 = 2 dt^2$ contribution. This is absent in the MA approximation, where different propagators always carry different momenta. It follows that the MA second order self-energy diagrams capture 6 out of the 7 finite $O(t^2)$ contributions correctly. Similar considerations apply for higher order $t$ powers and for higher order diagrams, differences between the MA and the exact self-energy diagrams coming only from terms involving free propagators carrying equal momenta in non-crossed diagrams. However, the error from such missed terms becomes smaller and smaller as one goes to higher order diagrams because the percentage of self-energy diagrams with one or more pairs of free propagators of equal momenta decreases exponentially.

We conclude that the MA approximation captures most of the $t$ dependence of each self-energy diagram, while summing over all diagrams. This analysis suggests that MA should be quite accurate for any finite $g$ and $t$ values. In Section 2.2.3 we reinforce this conclusion by analyzing the sum rules of the spectral weight.

### 2.2.2 Variational Interpretation of the MA Approximation

It is also instructive to point out an alternative explanation for the meaning of the MA approximation. This was originally pointed out in Refs. [78] and [79], and it was largely motivated by the failure of MA to predict the proper
2.2. The Meaning of the MA Approximation

Figure 2.2: A second order diagram contribution to \( G(i, j, \omega) \).

location for the polaron+one-phonon continuum. This continuum arises from states that have a phonon excited very far from where the polaron is. As a result, their interactions are negligible and the energy of the system is simply the sum of the two. As we will show in Section 2.3, MA either predicts a wrong location for this continuum, or no continuum at all.

This alternative explanation for the meaning of MA comes from realizing that in real space, the meaning of the MA approximation is that it replaces all free-propagators \( G_0(i, j, \omega - n\Omega) \) appearing in all self-energy diagrams, by \( \delta_{i,j}\bar{g}_0(\omega - n\Omega) \). For example, consider the real-space, second-order Green’s function diagram depicted in Fig. 2.2. It has the exact value

\[
\sum_{i_1,j_1} G_0(i, i_1, \omega) \Sigma_{2,c}(i_1, j_1, \omega) G_0(j_1, j, \omega),
\]

(2.32)

where the contribution to self-energy from the second-order crossed self-energy diagram is:

\[
\Sigma_{2,c}(i_1, j_1, \omega) = g^4 G_0(i_1, j_1, \omega - \Omega) G_0(j_1, i_1, \omega - 2\Omega) G_0(i_1, j_1, \omega - \Omega).
\]

(2.33)

Within MA, \( \Sigma_{2,c}(i_1, j_1, \omega) \) is approximated as

\[
\Sigma_{2,c}^{(0)}(i_1, j_1, \omega) = \delta_{i_1,j_1} \Sigma_{2,c}^{(0)}(\omega) = \delta_{i_1,j_1} g^4 \bar{g}_0(\omega - \Omega) \bar{g}_0(\omega - 2\Omega) \bar{g}_0(\omega - \Omega).
\]

(2.34)

Inserting this into the Green’s function diagram removes one of the sums, and after Fourier transforming we find that the contribution of this diagram to \( G(k, \omega) \) is \( G_0(k, \omega) \Sigma_{2,c}^{(0)}(\omega) G_0(k, \omega) \). The same holds for all higher order diagrams. Summing all of them, we see that the free propagators in the proper self-energy parts have indeed been replaced by their momentum averages.

The reason why it is a good zero-order approximation to keep only the diagonal (in real space) contributions \( G_0(i, j, \omega - n\Omega) \rightarrow \delta_{i,j}\bar{g}_0(\omega - n\Omega) \) is
2.2. The Meaning of the MA Approximation

straightforward to understand, at least for low energies $\omega \sim E_0$, where $E_0$ is the GS energy. Because of interactions, $E_0 < -2dt$ (the polaron ground state is below the free electron continuum). It follows that for $\omega \sim E_0$, the free propagators $G_0(i, j, \omega - n\Omega)$ are needed at energies below the free electron continuum, and the larger $n$ is the further below the band-edge these energies are. However, it is well-known that for $\omega < -2dt$, the free propagator decreases exponentially with increasing distance $|i - j|$. The reason is that there are no free-electron eigenstates outside the free-electron band, and the electron has to tunnel from one site to another. For example, in one dimension [80]

$$G_0(i, j, \omega) = e^{ik_0|i-j|}G_0(i, i, \omega) = e^{ik_0|i-j|}\bar{g}_0(\omega),$$

(2.35)

where $k_0$ is the first quadrant solution of the equation $\omega = -2t\cos k_0$. For $\omega < -2t$, $k_0 = i\kappa$, where $\kappa > 0$ increases as $\omega$ decreases. We also used the fact that in terms of its Fourier transform:

$$G_0(i, i, \omega) = \frac{1}{N} \sum_k e^{ik(R_i - R_i)}G_0(k, \omega) = \bar{g}_0(\omega).$$

(2.36)

It follows that at low energies, $\text{MA}^{(0)}$ keeps the largest contributions to the self-energy (from real-space diagonal terms), while ignoring exponentially small contributions coming from off-diagonal terms. This is expected to become more and more accurate for higher order diagrams with many phonons: the larger $n$ is, the faster the exponential decay with distance of $G_0(i, j, \omega - n\Omega)$ becomes.

Based on these arguments one expects at least low-energy properties to be well described by MA, at least if $\Omega$ is not too small. Comparison with numerics in Section 2.3 will validate this claim.

Taking this interpretation of MA one step further, it is clear that MA is equivalent to a variational approach where the eigenfunctions are built within a restricted Hilbert space that allows phonons only at one site. Indeed, it can be verified straightforwardly that wave functions of general type

$$|\Phi\rangle = \sum_{n=0}^{\infty} \sum_{i,j} \alpha_n(i, j)c_i^{\dagger}(b_j^{\dagger})^n|0\rangle$$

(2.37)

lead to the same spectrum as MA. Of course, the highest probability is to find the electron and phonon cloud at the same site, however hopping of the electron away from the phonon cloud is needed to allow for polaron motion.

This observation immediately explains the absence of the polaron+one-phonon continuum, since within this restricted Hilbert space it is not allowed
to have a single phonon far from where the polaron is. To compensate for the missing continuum’s weight and to fulfill the sum rules discussed in the next section, the GS $q_p$ weight is increased within MA. This will be shown by numerics in Section 2.3, as well as in Fig. 3.1(b) of the next chapter.

### 2.2.3 Sum Rules for the MA Approximation

For an even better idea of the accuracy of the MA approximation, we consider the sum rules for the spectral weight $A(k, \omega) = -(1/\pi) \text{Im} G(k, \omega)$, defined by

$$M_n(k) = \int_{-\infty}^{\infty} d\omega \omega^n A(k, \omega).$$  \hspace{1cm} (2.38)

For a problem of this type (single dressed particle), the sum rules can be calculated exactly to arbitrary order. The usual approach is based on the equation of motion technique [28]. We review it briefly here in order to make a few useful observations. The key step is to rewrite $\omega = (i \frac{d}{d\tau})^n e^{-i\omega \tau} \bigg|_{\tau=0}$, so that we have:

$$M_n(k) = -\frac{1}{\pi} \text{Im} \left( i \frac{d}{d\tau} \right)^n \int_{-\infty}^{\infty} d\omega e^{-i\omega \tau} G(k, \omega) \bigg|_{\tau=0}.$$  \hspace{1cm} (2.39)

The integral is now simply $G(k, \tau \rightarrow 0^+)$. Using the definition of Eq. (1.13), we find, for any $\tau > 0$:

$$\left( i \frac{d}{d\tau} \right)^n G(k, \tau) = -i \Theta(\tau) \langle 0 | c_k \mathcal{H}^n e^{-i\mathcal{H}\tau} c_k^\dagger | 0 \rangle,$$  \hspace{1cm} (2.40)

so that the sum rules simplify to:

$$M_n(k) = \langle 0 | c_k \mathcal{H}^n c_k^\dagger | 0 \rangle.$$  \hspace{1cm} (2.41)

These vacuum expectation values can be evaluated directly with some effort. We find

$$M_0(k) = 1,$$

$$M_1(k) = \varepsilon_k,$$

$$M_2(k) = \varepsilon_k^2 + g^2,$$

$$M_3(k) = \varepsilon_k^3 + 2g^2\varepsilon_k + g^2\Omega,$$  \hspace{1cm} (2.42)

and so on.

One very important conclusion that can be drawn from this derivation is that these sum rules have the same functional dependence on the energy...
scales \( t, \Omega, g \) anywhere in the parameter space. Of course, in various asymptotic regimes, different terms dominate the overall value (e.g. \( M_2(k) \approx g^2 \) if \( g \gg t \) while \( M_2(k) \approx \varepsilon_k^2 \) if \( g \ll t \)). However, this shows that if one can evaluate the sum rules exactly in any asymptotic regime, for instance by using perturbation theory, the results hold true everywhere in the parameter space, even where perturbation fails.

The second important conclusion one can draw from Eq. (2.41) is that each term in \( M_n(k) \) is proportional to \( t^p \Omega^m g^{n-m-p} \), where \( 0 \leq p, m, n-m-p \leq n \) are integers, i.e. the sum rule \( M_n \) is a polynomial of total order \( n \) in the energy scales of the problem. More complicated dependence on \( t, \Omega, g \), for example through \( \exp(-g^2/\Omega^2) \), simply cannot appear (see below).

The first conclusion suggests an alternative derivation of the sum rules, which can also be used for the MA and SCBA sum rules. Namely, we use the diagrammatic perturbational expansion of the Green’s function valid for \( g \ll t \) to evaluate directly the integrals \( \int_{-\infty}^{\infty} d\omega \omega^n G(k, \omega) \) and retain the imaginary part. In this case, \( G(k, \omega) = \sum_{n=0}^{\infty} \sum_{i=1}^{s_n} D_{n,i}(k, \omega) \), where \( D_{n,i}(k, \omega) \) are all Green’s function diagrams of order \( n \), i.e. containing \( n \) phonon lines. The multiplicity \( s_n = (2n-1)!! = 1 \cdot 3 \cdots (2n-1) \). Each diagram \( D_{n,i}(k, \omega) \) is a product of \( 2n+1 \) free propagators, summed over internal phonon momenta. Since for large frequency each \( G_0(q, \omega) \to (\omega + i\eta)^{-1} \), it follows that for \( |\omega| \to \infty \), each \( D_{n,i}(k, \omega) \to g^{2n}/(\omega + i\eta)^{2n+1} \).

Since any integrand that decreases faster than \( 1/\omega^2 \) has a vanishing contribution to Eq. (2.38), it follows that the diagrams of order \( n \) only contribute to the sum rules \( M_p(k) \) with \( p \geq 2n \). Thus, even though \( G(k, \omega) \) is the sum of an infinite number of diagrams, only a finite number of them, of low order, contribute to any given sum rule and the calculation can be done. The same holds true for the MA sum rules, the only difference being that the self-energy parts in the Green’s functions diagrams are replaced with the corresponding MA self-energy parts.

Let us analyze the differences between contributions of the exact and of the MA diagrams to the sum rules. It is straightforward to verify that:

\[
- \frac{1}{\pi} \text{Im} \int_{-\infty}^{\infty} d\omega \omega^{2n} g^{2n} \prod_{i=1}^{2n+1} G_0(q_i, \omega - \Omega_i) = g^{2n}, \tag{2.43}
\]

\[
- \frac{1}{\pi} \text{Im} \int_{-\infty}^{\infty} d\omega \omega^{2n+1} g^{2n} \prod_{i=1}^{2n+1} G_0(q_i, \omega - \Omega_i) = g^{2n} \sum_{i=1}^{2n+1} (\Omega_i + \varepsilon_{q_i}), \tag{2.44}
\]
2.2. The Meaning of the MA Approximation

and

\[-\frac{1}{\pi} \text{Im} \int_{-\infty}^{\infty} d\omega \omega^{2n+2} g^{2n} \prod_{i=1}^{2n+1} G_0(q_i, \omega - \Omega_i) \]

\[= g^{2n} \left[ \sum_{i=1}^{2n+1} (\Omega_i + \epsilon_{q_i})^2 + \sum_{i<j} (\Omega_i + \epsilon_{q_i})(\Omega_j + \epsilon_{q_j}) \right]. \quad (2.45)\]

Both the exact and the MA diagrams of order \(n\) contain products of the general form \(g^n \prod_{i=1}^{2n+1} G_0(q_i, \omega - \Omega_i).\) Some of the free propagators are actually \(G_0(k, \omega)\) (always the first and the last one, but there can also be intermediary ones connecting proper self-energy parts). All other free propagators have momenta dependent on the phonon momenta, which are summed over (in the exact diagrams), or are individually averaged over (in the MA diagrams). Since there is one-to-one correspondence between the number of exact vs. MA diagrams and their topologies, and since \(\sum q \epsilon_q = 0,\) it follows that the exact and the MA diagrams of order \(n\) give precisely the same contributions to \(M_{2n}(k)\) and \(M_{2n+1}(k).\) Differences appear in the contribution to \(M_{2n+2}(k)\) if there is at least one pair of propagators in any of the self-energy parts of the exact diagram that carry the same momenta. In this case, the corresponding \(\epsilon_{q_i} \epsilon_{q_j}\) averages to \(2d^2\) when the sums over phonon momenta are carried for the exact diagrams, whereas these terms always average to zero for the MA diagrams. Since most free propagators in the self-energy parts have different momenta, such differences are quantitatively small. This is especially true for large phonon frequencies \(\Omega,\) where the contributions proportional to \(\Omega\) captured correctly by MA scale like some power of \(2n + 1.\) This analysis can be continued for higher sum rules, with similar conclusions.

We can now summarize our findings. Only the 0th order Green’s function diagram [the free propagator \(G_0(k, \omega)\)] contributes to \(M_0(k)\) and \(M_1(k).\) Since this is included correctly in the MA and the SCBA cases, both give the correct \(M_0(k)\) and \(M_1(k).\) In fact, this diagram contributes an \(\epsilon_k^n\) to \(M_n(k),\) which is always the leading power in \(t\) contribution. The 1st order diagram is also exact in both the MA and SCBA cases, therefore both give the correct \(M_2(k)\) and \(M_3(k)\) sum rules as well. Differences appear from \(M_4\) onwards. Because SCBA only keeps 2 out of the 3 second order Green’s function diagrams, and 5 out of 15 third order diagrams, etc., the leading

\(^1\)Here we have assumed nearest-neighbour tight-binding. Such terms do not vanish for next-nearest neighbour hopping, for example, but we do not expect such terms to have a significant contribution to the spectral weight sum rules.
2.2. The Meaning of the MA Approximation

![Figure 2.3: Ratio of MA (squares), respectively SCBA (circles) sum rules and the exact sum rules, vs. order $n$. Results are for 1D and $k = 0$, $\Omega = 0.5t$ and $\lambda = 0.5, 1, 2$.](image)

Terms in $g$ are $2g^4$ instead of $3g^4$ in $M_4(k)$, $5g^6$ instead of $15g^6$ in $M_6(k)$, etc. This shows that SCBA fails badly all sum rules with $n > 3$ in the strong coupling regime where the term proportional to $g^{2n}$ gives the most significant contribution to $M_{2n}(k)$ (similar conclusions hold for odd sum rules). So even though SCBA always satisfies exactly the first 4 sum rules, it is a bad approximation for large $g$, where big discrepancies appear for $n > 3$.

MA satisfies $M_4(k)$ and $M_5(k)$ exactly as well, because these only depend on having the correct number and topology for the 2nd order diagrams. MA fails from $M_6(k)$ onwards, however in a very different manner than SCBA. The leading term in $g^6$ has the correct pre-factor, because MA has the correct number of 3rd order diagrams. The error comes from the 2nd order diagram containing the non-crossed self-energy diagram, as discussed.
2.2. The Meaning of the MA Approximation

Figure 2.4: Ratio of MA (squares), respectively SCBA (circles) sum rules and the exact sum rules, vs. order $n$. Results are for 1D and $k = \pi$, $\Omega = 0.5t$ and $\lambda = 0.5, 1, 2$. 
2.2. The Meaning of the MA Approximation

Indeed, instead of the exact sum rule:

\[
M_6(k) = \varepsilon_k^6 + g^2 \left[ 5\varepsilon_k^4 + 6t^2(2d^2 - d) + 4\varepsilon_k^3 \Omega + 3\varepsilon_k^2 \Omega^2 \\
+ 6dt^2(\varepsilon_k^2 + \varepsilon_k \Omega + 2\Omega^2) + 2\varepsilon_k \Omega^3 + \Omega^4 \right] \\
+ g^4(18dt^2 + 12\varepsilon_k^2 + 22\varepsilon_k \Omega + 25\Omega^2) + 15g^6,
\]

(2.46)

MA finds a sum rule \( M_{6}^{\text{MA}}(k) = M_6(k) - 2dt^2g^4 \). The leading terms in the \( g \ll t \) and \( g \gg t \) limits are always exact (as expected, since MA becomes exact in these limits), and this is true for all orders \( n \). For \( n \geq 6 \) some of the cross terms are missing, but these are a minority related to non-crossed diagrams, as explained. We therefore expect the MA sum rules to remain highly accurate for larger \( n \) values. That this is indeed true for higher sum rules is shown numerically in Figs. 2.3 and 2.4, where we plot the ratio of the MA respectively SCBA sum rules, and the exact sum rules of same order \( n \).

The results shown are for 1D and \( k = 0 \) and \( \pi \), but similar trends are found in the other cases. For \( k = 0 \) all the spectral weight is at negative frequencies, therefore \( M_n(0) \) alternate signs for even/odd \( n \), and this is reflected in the non-monotonic behavior with \( n \). For \( k = \pi \), most of the weight is at positive frequencies and sum rules are always positive. The magnitude of the exact sum rules increases roughly exponentially with \( n \), for instance for \( \lambda = 2 \) and \( \Omega = 0.5t \), \( M_{14}(0) = 119 516 000 \). For \( k = 0 \) and \( \lambda = 0.5 \), both MA and SCBA are reasonably accurate, with a slight edge for MA at higher \( n \). However, MA is clearly much more accurate for all the other cases shown, and its accuracy is expected to improve even more as one moves further into the asymptotic regions of weak or strong coupling.

Before ending this section, one more issue needs to be addressed. It is obvious that the MA sum rules must capture correctly the contributions to \( M_n(k) \) proportional to \( t^n \) and to \( g^n \), since MA is exact for both \( g = 0 \) and \( t = 0 \). One may assume that this alone is sufficient for a good interpolation at finite \( t \) and \( g \). That this is not so is shown by the generalized LF approximation, which is also exact for \( t = 0 \) or \( g = 0 \). However, in this approximation one finds \( M_n^{\text{LF}}(k) = 1, M_1^{\text{LF}}(k) = \varepsilon_k \exp \left( -\frac{g^2}{12} \right) \), etc. \( M_1 \) and all higher sum rules have unacceptable dependence on the energy scales \( g \) and \( \Omega \) (see second observation above), even though they become exact asymptotically. As shown in the next section, this approximation indeed performs rather poorly for finite \( t \) and \( g \).

We conclude that while MA satisfies exactly the first 6 sum rules, it remains accurate for all higher order sum rules, and is asymptotically exact. This is another argument in favor of the accuracy of this approximation.
over the whole parameter space. In the next section, we compare the MA predictions to those of existing numerical simulations to further support this claim.

2.3 Results

We first list the explicit expressions of the momentum averaged Green’s function $\bar{g}_0(\omega)$ of Eq. (2.22). For nearest-neighbour hopping, these are straightforward to derive:

\[ \bar{g}^{1D}_0(\omega) = \frac{\text{sgn}(\omega)}{\sqrt{(\omega + i\eta)^2 - 4t^2}}, \quad (2.47) \]

\[ \bar{g}^{2D}_0(\omega) = \frac{2}{\pi(\omega + i\eta)} K\left(\frac{4t}{\omega + i\eta}\right), \quad (2.48) \]

and

\[ \bar{g}^{3D}_0(\omega) = \frac{1}{2\pi^2 t} \int_0^{\pi} dk_z \text{sgn}\nu |\nu| K(\nu) \quad (2.49) \]

respectively, where

\[ \nu = \frac{4t}{\omega + 2t \cos(k_z a) + i\eta}, \quad (2.50) \]

and

\[ K(\nu) = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - \nu^2 \cos^2 \phi}} \quad (2.51) \]

is the complete elliptical function of the first kind [81]. These integrals can be performed numerically very efficiently. More generally, for any free electron dispersion $\epsilon_k$ to which corresponds the free electron density of states (DOS) $\rho_0(\epsilon) = (1/N) \sum_k \delta(\epsilon - \epsilon_k)$, we have $\bar{g}_0(\omega) = \int_{-\infty}^{\infty} d\epsilon \rho_0(\epsilon)(\omega - \epsilon + i\eta)^{-1}$ (also see Appendix A).

In our calculations we employ a small but finite value for $\eta$. This moves the poles of the Green’s function off of the real axis and changes the $\delta$-peaks of the spectral weight $A(k, \omega) = \sum_\alpha |\langle \alpha | c_k^\dagger | 0 \rangle|^2 \delta(\omega - E_\alpha)$ into Lorentzians. In practice, it is necessary to choose $\eta$ small enough to allow detection of the Lorentzian peaks in regimes where the $qp$ weight $Z$ is extremely small. This is relevant when ground state and low-energy properties are investigated. In such cases, we typically use $\eta/t \approx 10^{-5}$. For cases where $A(k, \omega)$ is plotted for a large frequency range, we use $\eta$ large enough to allow detection of the Lorentzian peaks with modest resolution in the step $\delta\omega$. The values used in these cases are listed explicitly.
2.3. Results

The self-energy $\Sigma^{(0)}_{\text{MA}}(\omega)$ is then calculated easily from Eq. (2.25) by truncating the continued fraction to a high-enough level. For an error of order $\epsilon$ it is necessary to go to a level with $n$ such that $ng_0^2(\omega - n\Omega)\bar{g}_0(\omega - (n + 1)\Omega) \leq \epsilon$. Using the fact that for large enough $n$ we can approximate $\bar{g}_0(\omega - n\Omega) \approx \bar{g}_0(\omega - (n + 1)\Omega) \approx -1/(n\Omega)$, it follows that we must have

$$n > \frac{1}{\epsilon} \frac{g^2}{\Omega^2}.$$  \hspace{1cm} (2.52)

This result is sensible because $g^2/\Omega^2$ is the average number of phonons in the polaron cloud as calculated from strong coupling perturbation theory [see Eq. (1.36)]. That is, in order to evaluate $\Sigma^{(0)}_{\text{MA}}$ with decent accuracy we are required to include all diagrams up to those diagrams containing the average number of phonons in the polaron cloud. A condition similar to Eq. (2.52) is obtained for the SCBA self-energies. In practice we always take $N$ to be much larger than in Eq. (2.52) and verify that doubling $N$ gives a change no more than a threshold much smaller than our infinitesimal parameter $\eta$. Thus, the error in all data points shown in this chapter is less than the thickness of the lines/symbols used for the plots. With an explicit form for $\Sigma^{(0)}_{\text{MA}}(\omega)$ we are now in a position to calculate the Green’s function $G^{(0)}_{\text{MA}}(0, k, \omega)$ and extract various polaron properties.

2.3.1 Polaron Ground State Properties

We begin by discussing polaron GS properties. Most of these are already known from numerical studies, but they give us an opportunity to further test the accuracy of the MA approximation. Given the simplicity and efficiency of the MA approximation, we can also investigate higher dimensionality and larger parameter ranges than typical numerically intensive approaches. In this section, we use for comparison 1D and 2D numerical results obtained from diagrammatic Quantum Monte Carlo (QMC) simulations from Ref. [26] unless otherwise noted.

For $k = 0$, we track the energy and weight of the lowest pole in the spectral weight, which give the ground state energy $E_0$ and the ground state quasi-particle weight $Z_0 = |\langle GS|c^\dagger_{k=0}|0\rangle|^2$. Using the Hellmann-Feynman theorem [82], we then find the average number of phonons in the ground state to be

$$N_{\text{ph}} \equiv \langle GS| \sum_q b^\dagger_q b_q |GS\rangle = \frac{\partial E_0}{\partial \Omega}. \hspace{1cm} (2.53)$$
2.3. Results

Note that one can also calculate the correlation function:

\[ \langle GS | \sum_i c_i^\dagger c_i (b_i^\dagger + b_i) | GS \rangle = \frac{\partial E_0}{\partial g} \]  

(2.54)

just as easily. We do not show it here because we do not have the corresponding numerical data for the comparison, however some typical results for this quantity are shown at the end of this section. Also note that all these quantities can be calculated similarly for other eigenstates. We will show such results in other sections.

We also show the effective mass, \( m^* \). Because the MA self-energy is momentum independent for this simple Holstein model, one has [21]:

\[ \frac{m^*}{m} = \frac{1}{Z_0} = 1 - \frac{d\Sigma_{MA(0)}(\omega)}{d\omega} \bigg|_{\omega=E_0}. \]  

(2.55)

This result also gives us a consistency check on our calculations. For MA we generally show effective mass results obtained from the first equality.

A comparison of the 1D results for these four quantities as obtained with QMC (black circles) and the different approximations is shown in Figs. 2.5 and 2.6. As expected, SCBA (blue line) fares well at small couplings \( \lambda \) but very poorly at strong couplings. The generalized LF (green line) is asymptotically exact, but quite wrong at finite \( \lambda \). Because these are GS properties, we can also compare to the perturbation theory results in both asymptotic limits. For weak couplings we use the Rayleigh-Schrödinger perturbation theory result [RS, violet line, Eq. (1.22)], and for strong couplings we use the second order perturbational theory result [PT, cyan line, Eq. (1.33)]. The GS energy is simply \( E_0 = E_{k=0} \). \( N_{ph} \) is obtained from \( E_0 \) as before; \( m^* \) can be evaluated from the second derivative of \( E_k \) with respect to \( k \), and \( Z_0 \) is extracted from the effective mass [Eq. (2.55)].

Fig. 2.5 shows that one or the other of these perturbational values describe the GS energy \( E_0 \) quite well, especially for the larger \( \Omega \) value. However, the agreement for the other quantities is somewhat poorer, especially at strong couplings (PT and LF give identical results for \( Z_0 \)). Part of the reason is that the \( t^2 \) term in Eq. (1.33) has in fact a more complicated dependence of \( g \) and \( \Omega \), which is only asymptotically equal to the one used here [31]. More importantly, neither perturbational theory describes well the crossover regime, or can be easily applied to high energy states.

Clearly, MA (red line with red symbols, for easier comparison with QMC data) has the best agreement with the QMC data. As expected from the sum rule analysis and the discussion on the convergence of \( \Sigma_{MA(0)} \), MA improves for larger \( \Omega \). The worst disagreements we ever found are the ones
2.3. Results

Figure 2.5: Ground state results in 1D. Shown as a function of the coupling $\lambda$ are the ground state energy $E_0$ and the qp weight $Z_0$. The left panels correspond to $\Omega/t = 0.1$ and the right ones to $\Omega/t = 0.5$. PT and LF give identical results for the qp weight.
Figure 2.6: Ground state results in 1D. Shown as a function of the coupling \( \lambda \) are the average number of phonons \( N_{\text{ph}} \) and the effective mass \( m^* \) on a logarithmic scale. The left panels correspond to \( \Omega/t = 0.1 \) and the right ones to \( \Omega/t = 0.5 \).
shown in the intermediary \( \lambda \) regime for \( \Omega/t = 0.1 \). Even there, the error in the GS energy is always below 5%. The \( qp \) weight has a more significant disagreement, however note that it indeed becomes asymptotically correct for \( \lambda < 0.2 \) and \( \lambda > 0.8 \). The second claim is supported by the \( m^* \) data, which indeed shows convergence towards the expected PT values. Most importantly, even though it is quantitatively somewhat wrong in this intermediary regime for small \( \Omega \), the MA approximation is the only one that clearly captures the crossover from the large to the small polaron, which is accompanied by the collapse of the \( qp \) weight and the increase in the number of phonons trapped in the cloud, and thus of the effective polaron mass.

The comparison with the QMC data for 2D polarons is shown in Figs. 2.7 and 2.8. Here MA gives excellent agreement at all couplings \( \lambda \). (Some of this is because the results correspond to larger \( \Omega/(4dt) \) values. We do not show \( \Omega/t = 0.2 \) results because we lack QMC data for comparison. Given the sum rule arguments, we expect the agreement to be better than for the 1D, \( \Omega/t = 0.1 \) case). The agreement is all the more remarkable when considering what a numerically trivial task it is to evaluate the MA results compared to the QMC simulations. The physics is similar to that seen in 1D, however the crossover from the large (light) polaron at weak couplings, to the small (heavy) polaron at strong couplings, becomes somewhat sharper, especially for smaller \( \Omega/t \) values.

We also show MA results in 3D, see Figs. 2.9 and 2.10. There are very few numerical three-dimensional results available for the Holstein polaron, due to the computational effort required to investigate such cases. Ground state properties of the three-dimensional polaron have been calculated using QMC [46] and variational methods [59] for a single phonon frequency of \( \Omega/t = 1.0 \). Comparison of the MA ground state energy and effective mass to those found in Ref. [46] shows excellent agreement. The good agreement with one or the other perturbational theories for most coupling strengths further supports the accuracy of MA in 3D. The crossover from large to small polaron becomes even sharper, especially for lower \( \Omega \). It is still located in the neighbourhood of \( \lambda \approx 1 \). This influence of dimensionality on the crossover regime of the Holstein polaron has also been reported by Ku et. al in Ref. [59].

Using the Hellmann-Feynman theorem like in Eqs. (2.53) and (2.54) also allows us to separate the individual contributions of

\[
T_{elec} = \langle GS | \sum_k \varepsilon_k c_k^\dagger c_k | GS \rangle,
\]

(2.56)
Figure 2.7: Ground state results in 2D. Shown as a function of the coupling \( \lambda \) are the ground state energy \( E_0 \) and the \( qp \) weight \( Z_0 \). The left panels correspond to \( \Omega/t = 0.5 \) and the right ones to \( \Omega/t = 1.0 \). PT and LF give identical results for the \( qp \) weight.
2.3. Results

Figure 2.8: Ground state results in 2D. Shown as a function of the coupling $\lambda$ are the average number of phonons $N_{ph}$ and the effective mass $m^*$ on a logarithmic scale. The left panels correspond to $\Omega/t = 0.5$ and the right ones to $\Omega/t = 1.0$. 

46
2.3. Results

Figure 2.9: Ground state results in 3D. Shown as a function of the coupling $\lambda$ are the ground state energy $E_0$ and the qp weight $Z_0$. The left panels correspond to $\Omega/t = 0.5$ and the right ones to $\Omega/t = 1.0$. The QMC results are from Ref. [46]. PT and LF give identical results for the qp weight.
2.3. Results

Figure 2.10: Ground state results in 3D. Shown as a function of the coupling \( \lambda \) are the average number of phonons \( N_{\text{ph}} \) and the effective mass \( m^* \) on a logarithmic scale. The left panels correspond to \( \Omega/t = 0.5 \) and the right ones to \( \Omega/t = 1.0 \). The QMC results are from Ref. [46].
2.3. Results

\[ E_{\text{ph}} = \langle GS | \Omega \sum_q b_q^\dagger b_q | GS \rangle = \Omega N_{\text{ph}}, \quad (2.57) \]

and

\[ V_{\text{corr}} = \langle GS | g \sum_i c_i^\dagger c_i \left( b_i^\dagger + b_i \right) | GS \rangle \quad (2.58) \]

to the total GS energy. Plots of these individual contributions as a function of coupling strength \( \lambda \) are shown in Fig. 2.11 for \( d = 1, 2, 3 \). The results for the kinetic energy of the electron are in good agreement with those found in Ref. [49] for \( d = 1, 2, 3 \), using QMC cluster sizes of 32, 12, and 6, respectively. We expect that the agreement for \( d = 2, 3 \) would be further improved with larger cluster sizes, however, this is computationally prohibitive for QMC simulations [49]. As expected, the kinetic energy is close to \(-2dt\) at weak couplings, but it becomes vanishingly small in the strong coupling limit, where the polaron becomes very heavy. The phonon energy \( E_{\text{ph}} \) increases roughly like \( g^2/\Omega \) in the strong coupling limit, where \( N_{\text{ph}} \approx g^2/\Omega^2 \). It follows that the decrease in the total GS energy is due to the interaction term, as expected. Note that this energy is proportional to the correlation of Eq. (2.54). Since \( E_0 \approx -g^2/\Omega \) in the strong coupling limit (see agreement with PT results), it follows that this correlation becomes asymptotically equal to \(-2g/\Omega\) in the strong coupling limit.

2.3.2 Low Energy States: Momentum Dependence

We can also calculate the same properties for the lowest energy state corresponding to each given momentum \( k \neq 0 \), to find the low-energy behavior of the polarons. In this section we present comparisons with available QMC results [26] for 1D and 2D systems, and with variational results for the 3D systems [59].

In Figs. 2.12 and 2.13 we show 1D results for the polaron dispersion \( E_k \), the associated \( qp \) weight \( Z_k \) and average phonon number \( N_{\text{ph}}(k) \) for two couplings. For the very weak \( \lambda = 0.25 \), we see that MA and SCBA are equally good at small \( k \), however for large \( k \) MA largely overestimates the energy. This observation actually reveals the key limitation of the MA approximation mentioned in Section 2.2.2: its failure to predict the polaron+one-phonon continuum which is expected to appear at a distance \( \Omega \) above the GS energy \( E_0 \). We will further discuss this discrepancy in Section 2.3.3.

As expected, the \( qp \) weight is large for small \( k \), where the main contribution to the eigenstate comes from the free electron state \( c_k^\dagger |0\rangle \). \( Z_k \) goes to zero for larger \( k \), since these are primarily linear combinations of states of
2.3. Results

Figure 2.11: GS expectation values for the electron kinetic energy (violet), the phonon energy (cyan), and the el-ph interaction (orange), as a function of $\lambda$, for $\Omega = 0.5t$ and $d = 1, 2, 3$. 
2.3. Results

type $c^\dagger_{k-q} b^\dagger_q |0\rangle$, as confirmed also by the average phonon number of about unity. Note that the RS perturbation works well for small $k$. It however breaks down at a finite $k$ where $\varepsilon_k \approx -2t + \Omega$, i.e. the free electron dispersion crosses into the continuum of electron plus one phonon states. Here, RS predicts an unphysical peak in the polaron dispersion [denominator of Eq. (1.22)] and it fails for larger $k$.

The second coupling $\lambda = 1$ is roughly in the crossover regime, see also Fig. 2.5. Here MA gives a much better agreement with QMC than SCBA or the perturbational theories. The polaron bandwidth is already renormalized and well below the weak coupling value of $\Omega$. In fact, as we show later, there is another bound state between these states and the polaron+one-phonon continuum. This is not captured in SCBA, which always predicts a polaron bandwidth of $\Omega$ with a continuum above, and roughly between zero and one average number of phonons as $k$ increases from 0 to $\pi$. For the strong coupling $\lambda = 1.96$, low-energy properties become almost $k$-independent, as expected since the Lang-Firsov impurity limit is being approached.

A second such comparison is possible for the 2D case with $\Omega = 0.5t$ and $\lambda = 0.845$, where QMC results are available, see Figs. 2.14 and 2.15. This coupling is on the weak side of the crossover, with the GS $qp$ weight still large, $Z_0 \approx 0.6$. Here MA is already doing better than SCBA. As in the 1D weak-coupling case, one can again see that the MA polaron bandwidth is slightly larger than $\Omega$: MA also somewhat overestimates the average number of phonons for large $k$ values. Overall, given that all of these results are in the crossover regime where the MA is at its worst, one can conclude that MA is also reasonably accurate in capturing low-energy polaron behavior.

In the 3D case, momentum-dependent variational results are available for energies and $qp$ weights, when $\Omega = 2.0t$ and $\lambda = 1.020833$. Although this coupling is in the center of the crossover regime, MA is still in excellent agreement with the variational results for all momenta $k$ along the high-symmetry cuts of the Brillouin zone, as shown in Fig. 2.16. The agreement is particularly good in this case because MA is highly accurate for large phonon energies.

2.3.3 Higher Energy States

The main motivation in trying to find an approximation for the Green’s function is that this quantity gives not only low-energy information, but the whole spectrum. We now compare MA predictions with various high-energy results available in the literature. Unfortunately there are much fewer of these, because the computational effort to obtain the whole spectrum
Figure 2.12: Lowest polaron eigenenergy for a given \( k \), \( E_k \), and the corresponding quasiparticle weight \( Z_k \). Results are for 1D, \( \Omega/t = 0.5 \), \( \lambda = 0.25 \) and 1. Only half of the BZ is shown.
2.3. Results

Figure 2.13: The average number of phonons in the polaronic cloud for the lowest polaron eigenenergy, \( N_{\text{ph}}(k) \) as a function of momentum in 1D. Results are for \( \Omega/t = 0.5, \lambda = 0.25, 1.0, 1.976 \). Only half of the BZ is shown.
Figure 2.14: Lowest polaron energy for a given $k$, $E_k$, and the corresponding $q_p$ weight $Z_k$. Results are for 2D, $\Omega/t = 0.5$, $\lambda = 0.845$. Three high-symmetry cuts in the Brillouin zone are shown.
2.3. Results

Figure 2.15: The average number of phonons in the polaronic cloud for the lowest polaron eigenenergy, $N_{ph}(k)$ as a function of momentum in 2D. Results are for $\Omega/t = 0.5$, $\lambda = 0.845$. Three high-symmetry cuts in the BZ are shown.
2.3. Results

Figure 2.16: Lowest polaron energy for a given \( k \), \( E_k \), and the corresponding \( qp \) weight \( Z_k \). Results are for 3D, \( \Omega/t = 2.0 \), \( \lambda = 1.020833 \). Three high-symmetry cuts in the Brillouin zone are shown. The variational results (VM) are from Ref. [59].
2.3. Results

through the usual numerical approaches is generally forbidding.

We begin with a comparison against exact diagonalization (ED) 1D data, from Ref. [53]. The results are shown in Fig. 2.17. As already discussed, for weak coupling there is a polaron+one-phonon continuum starting at \( E_0 + \Omega \), but for stronger coupling a second bound state appears below the continuum. In Fig. 2.17(a) we track the energy \( E_1 \) of the second \( k = 0 \) state. For small couplings, the results actually show the maximum DOS in the continuum, not its edge (the maximum is generally located close to the lower edge. This results again shows that MA somewhat overestimates this energy, which should be \( \approx \Omega \)). When \( E_1 < E_0 + \Omega \), there is a true discrete state. Note that panel (a) is in very good quantitative agreement with similar data shown in Fig. 8 of Ref. [53]. The only difference is for strong coupling, where the ED results show \( E_1 > E_0 + \Omega \) again, however, with significant finite-size dependence on the chosen Hilbert space cutoff.

We can thus find the coupling \( g/t \) where the second bound state appears, for different values of \( \Omega/t \). This line is shown in panel (b), together with the ED results. The agreement between the two data sets is excellent, even at small \( \Omega/t \) values where we expect MA to be less accurate. We also show in panel (c) the \( qp \) weight of this second bound state, where stable. These results are not given in Ref. [53], however one QMC point is available in Ref. [26], in good agreement with the MA prediction.

We now move to comparisons for the entire spectral weight \( A(k, \omega) \). In Figs. 2.18, 2.19 and 2.20 we show comparisons for a 1D system with \( \Omega = 0.4t \) and three different coupling strengths \( \lambda = 0.5, 1 \) and 2, respectively. In each case, data for 5 values of \( k \), namely 0, \( \pi/4 \), \( \pi/2 \), \( 3\pi/4 \) and \( \pi \) are shown. The numerical results (black line) are obtained using a variational method by De Filippis et al. [62]. Numerical results obtained through exact diagonalization of a finite system and by QMC, for the same parameters but somewhat different \( k \) values, are also presented by Hohenadler et al. in Refs. [50, 66]. These sets of numerical data are in good agreement with one another.

In all three cases the agreement between the MA and numerical results is very good. As expected, it is best for the largest \( \lambda \), but even for the smaller \( \lambda \) values, which are just below and within the crossover region, the agreement is very satisfactory. For \( \lambda = 0.5 \) and \( k = 0 \) (upper panel of Fig. 2.18) we see the polaron state as a Lorentzian peak (a broadening \( \eta = 0.1\Omega \) was used) which accounts for most of the weight, and a small continuum at a higher energy. MA overestimates the gap between the two, which should be \( \Omega \). As \( k \) increases, the polaron peak disperses but also loses significant weight, as discussed in the previous section. Most of the weight is now in the high
2.3. Results

Figure 2.17: (a) Energy gap $\Delta = E_1 - E_0$ between GS and first excited $k = 0$ state. A second bound state is stable if $\Delta < \Omega$ (here $\Omega = 0.5t$); (b) Line below which a second bound state appears: ED [53] (circles) and MA (squares); (c) MA $qp$ weight of the second bound state when stable (red squares), and that of the GS (black squares), for $\Omega/t = 0.5$. The circle is the one QMC result available for the $qp$ weight of the second bound state [26]. These results are for 1D.

energy continuum, located roughly near the corresponding $\varepsilon_k$ value. This simply shows that these higher energy states are not significantly affected by this rather weak coupling. The VM results show somewhat more structure in these continua than MA, but most of the weight occupies similar frequency ranges.

For $\lambda = 1$ and $k = 0$ (upper panel of Fig. 2.19), the MA result shows 3 Lorentzian peaks plus a continuum starting at $\omega/t = -1.6$. For the rather large $\eta$ used it is hard to distinguish which peaks come from individual poles, and which are true continua. This can be easily done by studying their behavior as a function of the broadening $\eta$, as shown in Fig. 2.21. The height of peaks corresponding to discrete states scales precisely like $1/\eta$, as
expected for Lorentzians. The continuum is affected very little by changes in \( \eta \), except the peak near its lower edge where the finite \( \eta \) smoothes out a singularity in the DOS. Since this singularity is not of the \( 1/\omega \) type, its scaling with \( \eta \) is different from that of the Lorentzians. The two lower states are closer to one another than \( \Omega \), however the MA results show no sign of the continuum that is expected to start at \( E_0 + \Omega \). Note that the numerical data in Fig. 2.19 shows more structure, that could be consistent with this continuum. We will address the issue of this continuum below. As \( k \) is increased (see Fig. 2.19) the low-energy peaks show some dispersion, but with a strongly renormalized bandwidth. At higher \( k \) most weight shifts again at high energies, in a rather broad continuum.

Finally, for \( \lambda = 2 \) and \( k = 0 \), Fig. 2.20 shows even more discrete peaks spaced by \( \Omega \). The GS is at \( E_0 \approx -4.25t \), but its weight is so small that it cannot be seen on this scale, unless \( \eta \) is decreased significantly. A continuum
2.3. Results

is seen above $\omega = -1.6t$. As $k$ is now increased, there is almost no dispersion of the discrete peaks, however the weight shifts again to an even broader high-energy continuum.

The issue of the continuum at $E_0 + \Omega$ in the exact case, and of its absence in the MA approximation for moderate and large couplings can be understood from Fig. 2.22. Here we show a comparison of $A(k, \omega)$ in 2D obtained from exact diagonalization [83] vs. MA results, for various couplings $\lambda$. The first remark is that MA captures quite well all the large-weight features, both as far as their energy and their weight are concerned. This is expected, given the good sum rules agreement demonstrated previously. However, the ED results clearly show more states than MA predicts. There is always a low-weight peak at precisely $E_0 + \Omega$ (the GS is marked by arrows for $\lambda > 1.125$). For couplings of up to $\lambda \approx 1$, this peak is followed by several nearby peaks with comparably low-weight, which can be argued to be part
2.3. Results

Figure 2.20: 1D spectral weight $A(k, \omega)$ vs. $\omega$, for $k = 0, \pi/4, \pi/2, 3\pi/4$ and $\pi$. MA results (red line) vs. data from Ref. [62] (black line). Parameters are $\Omega = 0.4t$, $\lambda = 2$, $\eta = 0.1\Omega$.

of the expected continuum. For larger $\lambda$, however, only the state at $E_0 + \Omega$ can still be seen, although more states suggesting more continua are seen between the large higher-energy peaks. The gradual disappearance of the first continuum is not surprising, since one expects its width to narrow exponentially as the coupling increases. Moreover, one expects that the largest contribution to this continuum is from states with one or more phonons, explaining their low $qp$ weight.

As far as MA is concerned, Fig. 2.22 suggests that for couplings where there is more than one discrete state, the very little weight in the $E_0 + \Omega$ and similar higher-energy continua is absorbed in the discrete states predicted by MA. This is consistent with the systematic up-shift of the MA peaks compared to the ED data.

In fact, it is straightforward to see that the MA approximation can only predict a continuum starting at $-2dt + \Omega$. A continuum is signaled by
2.3. Results

Figure 2.21: MA 1D spectral weight $A(k=0, \omega)$ vs. $\omega$, for $\Omega = 0.4t$, $\lambda = 1$, and $\eta/\Omega = 0.1, 0.05, 0.02, 0.01$. The first three peaks are discrete states (lorentzians) whereas the fourth marks the band-edge singularity of the continuum.

A finite imaginary part of $\Sigma_{\text{MA}}^{(\omega)}(\omega)$, and the lowest frequency where this can occur is that for which $\tilde{g}_0(\omega - \Omega)$ acquires a finite imaginary part [see Eq. (2.25)]. However, the imaginary part of $\tilde{g}_0(\omega)$ is proportional to the total density of states of the free model, i.e. it is finite for $\omega \in [-2dt, 2dt]$ for nearest-neighbour hopping. It follows that the MA continuum always starts precisely at $-2dt + \Omega$. This explains why for small $\lambda$, where there is only one peak below this continuum, the gap between the two is somewhat larger than the expected $\Omega$ value: the GS energy decreases below $-2dt + \Omega$ with increasing $\lambda$, whereas the continuum edge is pinned at $-2dt + \Omega$, in the MA approximation. As the coupling increases, bound states start to split from this continuum, and spectral weight is shifted to lower energies, in good agreement with the sum-rule predictions of the exact solution. These new bound states have to account for the (small) weight that is present in lower energy continua, in the exact solution, and this is precisely what Fig. 2.22 shows. Clearly, a self-energy that would account for these continua as well would have to be more complicated than that of Eq. (2.25).
Figure 2.22: 2D spectral weight $A(k = 0, \omega)$ for $\Omega = 0.5t$, $\eta = 0.01t$ and various $\lambda$ values, from exact diagonalization [83] (black) and MA (red). For $\lambda > 1.125$, the arrows point the GS location.

In the following, we use MA to investigate more properties of the Green’s function. We begin with Fig. 2.23, where we plot the $qp$ weight and average phonon numbers for a few of the higher-energy peaks, once they appear below the continuum. For comparison, we also show the already discussed GS results (black line). Results in higher dimension are qualitatively similar and we do not show them here. Unlike for the GS, both these quantities are non-monotonic functions of $\lambda$ for all higher-energy bound states. Each of these states disperses with $k$, like in Figs. 2.19 and 2.20 (more results for this are shown below), so an effective mass can be associated with each such band. This effective mass satisfies $m^*/m = 1/Z$, and therefore also shows non-monotonic behavior, first decreasing and then increasing as $\lambda$ is increased. The average phonon number in the $n$th state must approach $n + g^2/\Omega^2$ asymptotically, as can be verified in the Lang-Firsov limit. This is indeed observed in Fig. 2.23, however the plateaus seen at moderate $\lambda$ suggest some crossover from one to another type of wave function associated
2.3. Results

Figure 2.23: qp weight and average number of phonons in the GS (black line) and the next three higher \( n = 0 \) bound states, when they become stable according to MA. Results are for 1D, \( \Omega = 0.5t \).

with these higher levels. As \( \lambda \to \infty \), an infinite sequence of such bound states appear, as expected in the Lang-Firsov limit.

For a better illustration of the appearance of these bound states and of their evolution, we show contour plots of the spectral weight \( A(k, \omega) \). We begin by plotting \( A(k, \omega) \) as a function of \( k \) and \( \omega \), for fixed parameters \( g, t \) and \( \Omega \). In the left panels of Fig. 2.24 we show 1D results corresponding to \( \Omega = t \) and \( \lambda = 0.4, 1 \) and 2, respectively. Only half of the BZ is shown as \( G(k, \omega) = G(-k, \omega) \). Each of these MA contour plots takes below ten seconds to generate. Note that similar plots for the same parameters were provided by Hohenadler et al. in Ref. [66], based on a cluster perturbation theory approach. The agreement between the main features of our and their plots is excellent. As expected, their results do show a few more low-weight features at lower energies, below \( -2t + \Omega = -t \), in this case, where our continuum starts. Such contour plots are richer versions of plots like those shown in Figs. 2.18, 2.19 and 2.20. They illustrate basically the same points, although one can now also see clearly the dispersion of various features. For the lowest \( \lambda \), the free electron dispersion \( \varepsilon_k = -2t \cos(ka) \) is still almost visible, except that el-ph interactions split it into the lower
2.3. Results

polaron band, and the higher continuum. This continuum is not featureless, instead one can already see weight accumulating near its lower edge. As $\lambda$ increases, a new bound state will split off from it. This is seen for $\lambda = 1$, where there are 2 bound states below the continuum starting at $-1$ (for these parameters). The bandwidth of each of these states is now narrowed below $\Omega$. The weight in the continuum at higher energies is redistributed suggesting the impending formation of yet more bound states. Indeed, the $\lambda = 2$ results show 4 even narrower bound states below the continuum.

Similar behavior is expected, and indeed seen, in higher dimensions. Here we only show similar 2D contour plots, for $\Omega = 2t$ and $\lambda = 0.5, 0.945$ and 2, in the right panels of Fig. 2.24. The middle panel again agrees very well with data shown in Ref. [66]. In this case, the MA continuum starts at $-4t + \Omega = -2$. As $\lambda$ increases, we see again first 1, then 2 and then 4 bound states below the continuum. Their bandwidths decrease with increasing $\lambda$, so that the lowest band for $\lambda = 2$ is already almost dispersionless, even though its weight still varies with $k$. As in the 1D case, as the interaction becomes stronger, the weight in the continuum also redistributes itself, with strong resonances seen around multiples of the phonon frequency.

Another way to understand the dependence on the coupling $\lambda$ (or any other parameter) is to plot a contour of the spectral weight $A(k, \omega)$ vs. $\lambda$ and $\omega$ for a fixed value of $k$. Such a task is equally trivial at the MA level. In fact, one can also just as easily calculate and plot the total density of states, or spectral weight:

$$A(\omega) = \frac{1}{N} \sum_k A(k, \omega),$$

(2.59)

since within the MA approximation this is given by:

$$A(\omega) = -\frac{1}{\pi} \text{Im} \left[ \frac{1}{N} \sum_k \frac{\omega - \varepsilon_k - \Sigma_{\text{MA}}^{(0)}(\omega) + i\eta}{1} \right]$$

$$= -\frac{1}{\pi} \text{Im} \left\{ g_0 \left[ \omega - \Sigma_{\text{MA}}^{(0)}(\omega) \right] \right\}.$$

(2.60)

In Fig. 2.25 we show 4 such contour plots for the 1D case. The top left panel shows $A(k = 0, \omega)$ vs. $\omega$. For $\lambda = 0$ (non-interacting case), only one state exists at $-2t$, as expected. As the coupling turns on, the energy of this state (the ground state) decreases, but $k = 0$ weight is also transferred to higher energies, due to hybridization with the states in the electron-plus-one-phonon continuum. The MA continuum here starts at
2.3. Results

Figure 2.24: Contour plots of the 1D spectral weight $A(k, \omega)$ (left panels) and 2D spectral weight $A(k, \omega)$ (right panels) as a function of $k$ and $\omega$. The 1D parameters are $\Omega = t$, $\lambda = 0.4, 1$ and 2, respectively, and $\eta = 0.02$. These results are in excellent agreement with the results of Hohenadler et al., Ref. [66], and the GS results of Bonca et al., Ref. [53]. The 2D parameters are $\Omega = 2t$, $\lambda = 0.5, 0.945$ and 2, respectively, and $\eta = 0.02$. The 2D middle panel is in excellent agreement with the results of Hohenadler et al. given in Ref. [66].
2.3. Results

Figure 2.25: 1D results for $\Omega = 0.4t$ and $\eta = 0.04t$. The left panels show the spectral weight $A(k = 0, \omega)$ and the total spectral weight $A(\omega)$ vs. $\omega$ and $\lambda$. The right panels show the spectral weight $A(k = \pi, \omega)$ vs. $\omega$ and $\lambda$, on both a linear scale and a logarithmic scale.
2.3. Results

Figure 2.26: The left panels show 2D spectral results and the right panels show 3D spectral results, for $\Omega = 0.5t$ and $\eta = 0.04t$. The top panels show the spectral weight $A(k = 0, \omega)$ vs. $\omega$ and $\lambda$, and the bottom panels show the total spectral weight $A(\omega)$ vs. $\omega$ and $\lambda$. 
2.3. Results

For moderate and larger $\lambda$ one can clearly see how weight is rearranged inside the continuum as $\lambda$ increases, and new bound states split from it and move towards lower energies. The apparent “break” in the slope of the GS energy, as $\lambda$ increases, is now seen to occur when the first bound state approaches the GS, and is suggestive of an avoided crossing. From this point on the GS lowers its energy much faster, but its weight also decreases dramatically and it becomes difficult to see. The top right panel of Fig. 2.25 shows $A(k = \pi, \omega)$ vs. $\omega$, on a linear scale. At $\lambda = 0$, there is only one peak at $+2t$, as expected. As the coupling is turned on, this weight seems to spread around in a rather featureless, broad continuum. In fact, on a logarithmic scale (bottom right panel of Fig. 2.25) one can see that $k = \pi$ weight is pulled down into all the bound states, in agreement with the previous data we showed. This weight, however, is so small that it is not visible on the linear scale. Finally, the bottom left panel of Fig. 2.25 shows the total spectral weight, or DOS. At $\lambda = 0$ we see the usual 1D DOS, with the singularities near the band-edges rounded off because of the finite $\eta$ used. As $\lambda$ is turned on, one can recognize both the contribution from the $k = 0$ and $k = \pi$ states to the total DOS: each bound state has a finite bandwidth due to its dispersion (this is to be contrasted to the upper panels, where the bound states are true delta-functions, with a width defined by the broadening $\eta$). As the coupling strength increases, the number of bound states increases; they are spaced by roughly $\Omega$, their bandwidths narrow down, and their weights also decrease. In other words, they approach the expected Lang-Firsov behavior.

Thus, this figure actually answers the question posed in the Introduction of this chapter, regarding the evolution of the spectral weight from that of a free electron towards that of the impurity limit. While the MA results are certainly not exact, we can claim with a high degree of certainty that the main features are accurately captured, especially in the weak and in the strong coupling limits. This suffices to understand the physics of this problem, and given the simplicity of the approximation, to investigate in detail a number of other quantities we have not shown here, such as the self-energy. Of course, if one is interested in exact results for some particular set of parameters, numerical methods have to be used.

Qualitatively similar plots are obtained in 2D and 3D, as shown in Fig. 2.26. Of course, the $\lambda = 0$ DOS is very different, with a van-Hove singularity at $\omega = 0$ for the 2D case, and the characteristic nearest-neighbour hopping DOS in the 3D case. However, the appearance of multiple bands below the continuum as $\lambda$ increases and all the remaining phenomenology is very similar. Thus, we see no evidence of any qualitative differences in the polaron
2.4 Conclusion

In this chapter we analyzed the Green’s function of the Holstein polaron, using the \textit{momentum average} approximation, which consists of summing all the self-energy diagrams, but with each individually averaged over all its free propagator momenta. The resulting self-energy can be written as an infinite continuous fraction that is numerically trivial to evaluate. This procedure becomes exact in the limits $g = 0$ and $t = 0$.

We gauged the accuracy of this approximation by computing its corresponding spectral weight sum rules and comparing them against the exact sum rules, which are known for this type of Hamiltonian. We showed that the MA spectral weight satisfies exactly the first 6 sum rules. Even though this is quite impressive at first sight, it is actually no guarantee of overall accuracy, as the case of SCBA demonstrates. The SCBA spectral weight always satisfies exactly the first 4 sum rules, even at large couplings $\lambda$ where it predicts very wrong results. The meaningful criterion of accuracy for the sum rules is to show that the vast majority of terms in \textit{all sum rules}, and in particular the dominant terms in the various asymptotic limits, are captured by the approximation. MA indeed satisfies this very restrictive criterion.

The accuracy of the approximation was also tested by direct comparison with data obtained through numerically intensive methods. In all cases we obtain good agreement, especially considering the ease of evaluation of the MA results. The MA approximation is not exact, but in all cases the higher-weight features of the spectral weight are qualitatively well-described by the MA approximation. Trading some of the accuracy of numerically exact but often time-consuming methods in exchange for very fast results which capture the main features accurately is a useful approach when trying to understand the main aspects of the physics of a problem, as well as when one is concerned about comparison with experiments.

The key features not captured by the MA approximation are as follows. As discussed previously in this chapter, MA fails to predict the polaron+one-phonon continuum that must appear at $E_0 + \Omega$, where $E_0$ is the polaron ground state (GS) energy, and $\Omega$ is the frequency of the Einstein phonons. We noted that numerical simulations show that there is very little spectral weight in this continuum, hence its absence or wrong positioning does not significantly upset the agreement with the sum rules, but nevertheless, it would be reassuring to have an approximation that correctly predicts its
existence. The other shortcomings of the MA approximation are that it worsens as $\Omega \to 0$, and that the MA self-energy $\Sigma_{\text{MA}}(\omega)$ is independent of the momentum $k$ of the electron. Given how featureless the Holstein model is (el-ph coupling and phonon frequency are both constants) one may expect a rather weak momentum dependence of the self-energy, however it is certainly not entirely absent. We support this claim in Chapter 3, where we systematically improve MA for the Holstein model and find that it indeed has a weak momentum dependence. In Chapter 4 we take this a step further and show that the self-energies calculated using MA are explicitly momentum-dependent for more complicated Hamiltonians, and that the momentum-independent result found here is due to the simplicity of the Holstein model.

In conclusion, it is important to note that the existence and accuracy of approximations like MA is not guaranteed, in fact it can be regarded as a surprise for the case of the Holstein polaron, that has been under investigation for almost five decades. However, this demonstration of its existence and efficiency in the Holstein polaron problem gives some hope for making progress for the general class of strongly-correlated systems problems. One can always use some flavour of perturbation theory to understand behavior in asymptotic limits, but the really challenging problems are set in regimes where perturbation does not apply. The MA approximation suggests that one way to make non-trivial progress is to sum all diagrams, with each simplified enough so as to make the calculation feasible, but not so much as to really alter the physics. This is a very different approach from the usually employed summation of a subclass of diagrams. Note that there are many classes of problems with diagrams similar to the ones arising in the single polaron problem, although of course with different propagators and/or vertices. In the following chapter the shortcomings of this approximation will be addressed, particularly the absence of the el-ph continuum, and “cured” by systematically improving the MA approximation.
Chapter 3

Systematic Improvement of the MA Approximation

In the previous chapter we derived the momentum average (MA) approximation for the Green’s function of a single Holstein polaron. The essence of this approximation is to sum all diagrams contributing to the polaron self-energy, however each diagram is approximated to such a degree as to allow the analytical summation of the entire series. Specifically, each free propagator appearing in a self-energy diagram is replaced by its momentum average. The resulting MA self-energy is a trivial-to-evaluate continued fraction which gives remarkably accurate results over most of the parameter space, including intermediate electron-phonon coupling strengths where perturbational methods completely fail to capture the correct physics.

We also identified some of the reasons for this good agreement: first, the MA approximation becomes asymptotically exact for both very weak and very strong couplings. More importantly, the resulting MA spectral weight satisfies the first six spectral weight sum rules exactly, and remains highly accurate for all higher order sum rules. However, we also pointed out some shortcomings of the MA approximation: it fails to predict the polaron+one-phonon continuum, its accuracy worsens as $\Omega \to 0$, and it produces a momentum-independent self-energy.

In this chapter, we show how to systematically improve the MA approximation, generating a hierarchy of approximations that we call MA(\(n\)) (the original MA is MA(0) in this notation). As explained below, the idea is to systematically improve the accuracy of the “simplified” self-energy diagrams. The results become more and more accurate as \(n\) increases, for example, while the MA spectral weight satisfies only the first 6 sum rules exactly, this improves to 8 and 10 exact sum rules respectively for MA(1) and MA(2) spectral weights. While the numerical effort also increases, it is still trivial for the \(n = 1\) and \(n = 2\) levels that we discuss explicitly here. Level MA(1) already solves the continuum problem, while all levels with \(n \geq 2\) produce momentum-dependent self-energies. The accuracy in the limit $\Omega \to 0$ is also shown to improve significantly with increasing $n$. In
3.1 Higher Level MA

In our first description of the MA approximation we showed that the approximation amounts to simply replacing all free propagators in the self-energy expansion by their momentum average $\bar{g}_0(\omega)$. Through a detailed analysis of diagrammatics, spectral weight sum rules, and a comparison to available numerical data, we were then able to demonstrate the accuracy and range of validity for the approximation. In Section 2.2.2 we provided an alternative explanation for the good accuracy of the MA approximation. There it was shown that in the MA self-energy expansion we retain only the diagonal terms of the free propagator (in real space):

$$G_0(i, i, \omega - n\Omega) = \frac{1}{N} \sum_k e^{ik(R_i - R_i)}G_0(k, \omega - n\Omega) = \bar{g}_0(\omega - n\Omega).$$

(3.1)

Because the free propagator $G(i, j, \omega - n\Omega)$ decreases exponentially with distance $|i - j|$ for $\omega < -2dt$, and this exponential decay becomes faster with increasing $n$, we argued that this was a good approximation, at least for low energies.

The latter explanation suggests a systematic way towards improving the MA$^{(0)}$ approximation. The biggest error at low energies is due to the momentum average of the propagators of energy $\omega - \Omega$. This is the energy closest to the free electron continuum and therefore these propagators have the slowest decay in real space. The next slowest decay is for the propagators of energy $\omega - 2\Omega$, etc. If one could selectively keep these propagators exactly while momentum-averaging the ones with more phonons (lower energy, faster decay) this should improve the accuracy of the approximation at low energies. In fact, all sum rules would also be improved (individual
3.1. Higher Level MA

Diagrams are more accurate) therefore one would expect an improvement at all energies. This is precisely what the higher levels MA\(^{(n)}\), \(n \geq 1\), achieve.

We define MA\(^{(n)}\) as being the approximation where in all self-energy diagrams, all free propagators with energy \(\omega - m\Omega\), where \(m \leq n\), are kept exactly, while the ones with more than \(n\) phonons are momentum averaged.

For convenience we repeat the relevant equations of motion from Eqs. (2.8) and (2.9) here:

\[
G(k,\omega) = G_0(k,\omega) \left[1 + \frac{1}{N} \sum_{q_1} f_1(q_1)G(k,\omega)\right]
\]

(3.2)

and

\[
f_n(q_1,\ldots,q_n) = G_0(k - q_T,\omega - n\Omega)
\times \left[g^2 \sum_{i=1}^{n} f_{n-1}(\ldots,q_{i-1},q_{i+1},\ldots) + \frac{1}{N} \sum_{q_{n+1}} f_{n+1}(q_1,\ldots,q_{n+1})\right].
\]

(3.3)

In terms of these equations that need to be solved, systematically improving the MA approximation is straightforward, since a propagator of energy \(\omega - m\Omega\) appears only once, in the right-hand side pre-factor of the equation for \(f_m(q_1,\ldots,q_n)\). It follows that if we keep the first \(n\) of Eq. (3.3) as they are, and approximate the equations for \(f_{n+1}, f_{n+2}, \ldots\) by momentum-averaging the free propagator appearing in the right-hand side pre-factor \(G_0(k - q_T,\omega - n\Omega) \to \bar{g}_0(\omega - m\Omega)\) if \(m \geq n + 1\), we achieve our goal, provided that we can find the solution of the resulting infinite system of coupled equations.

We now derive explicitly the solutions for \(n = 1\) and \(n = 2\) levels.

### 3.1.1 MA\(^{(1)}\) Level

In this case, the equations to be solved are:

\[
\Sigma_{\text{MA}^{(1)}}(k,\omega) = \frac{1}{N} \sum_{q_1} f_1^{(1)}(q_1)
\]

(3.4)

where

\[
f_1^{(1)}(q_1) = G_0(k - q_1,\omega - \Omega) \left[g^2 + \frac{1}{N} \sum_{q_2} f_2^{(1)}(q_1,q_2)\right],
\]

(3.5)
and for all \( n \geq 2 \),
\[
    f_n^{(1)}(q_1, \ldots, q_n) = \bar{g}_0(\omega - n\Omega)

\times \left[ g^2 \sum_{i=1}^{n} f_{n-1}^{(1)}(\ldots, q_{i-1}, q_{i+1}, \ldots) + \frac{1}{N} \sum_{q_{n+1}} f_{n+1}^{(1)}(q_1, \ldots, q_{n+1}) \right].
\]

(3.6)

As before, the dependence on \( k, \omega \) is implicitly assumed everywhere. We use the upper labels (1) because these are the approximative solutions corresponding to MA\(^{(1)}\).

The detailed solution of this infinite set of recurrent equations is discussed in Appendix B, but the approach is as follows. We use Eqs. (3.6) to find how \((1/N) \sum_{q_2} f_2^{(1)}(q_1, q_2)\) depends on \( f_1^{(1)}(q_1) \). This is then used in Eq. (3.5) to solve for \( f_1^{(1)}(q_1) \), and \( \Sigma_{\text{MA}^{(1)}} \) then follows from Eq. (3.4). The end result is
\[
    \Sigma_{\text{MA}^{(1)}}(\omega) = g^2 \bar{g}_0(\tilde{\omega}) \left[ \frac{1}{1 - g^2 \bar{g}_0(\tilde{\omega}) [A_2(\omega) - A_1(\omega - \Omega)]} \right],
\]

(3.7)

where [see Eq. (2.25)]:
\[
    \tilde{\omega} = \omega - \Omega - g^2 A_1(\omega - \Omega) = \omega - \Omega - \Sigma_{\text{MA}^{(0)}}(\omega - \Omega).
\]

(3.8)

The continued fractions \( A_1(\omega - \Omega), A_2(\omega) \) are defined in Eq. (2.26). This expression is slightly more complicated than \( \Sigma_{\text{MA}^{(0)}}(\omega) \), since it involves two different continued fractions, however it is still very trivial to compute.

Note that based on this and other results derived above, we can now calculate the MA\(^{(1)}\) expressions for the generalized Green’s functions
\[
    f_n^{(1)}(k, q_1, \ldots, q_n, \omega) \sim F_n(k, q_1, \ldots, q_n, \omega).
\]

(3.9)

These will be more accurate than the values obtained within the MA\(^{(0)}\) approximation, where none of the \( f_n^{(0)} \) expressions had any momentum dependence. These generalized Green’s functions contain further information about the polaron, for example regarding the phonon statistics.

As can be seen from Eq. (3.7), the self-energy is still momentum independent at this level. However, it is clear that this is because the Holstein model is so featureless. If, for e.g. the coupling was dependent on the phonon momentum, then the first self-energy diagram
\[
    \Sigma^{(1)} = \frac{1}{N} \sum_{q} |g(q)|^2 G_0(k - q, \omega - \Omega)
\]

(3.10)
3.1. Higher Level MA

would be \( k \) dependent, and so would \( \Sigma_{\text{MA}}^{(1)} \) (this diagram is taken into account exactly at the MA\(^{(1)} \) level). Indeed, this will be verified in Chapter 4 when we look at the generalization of the MA approximation to systems with el-ph coupling dependent on the phonon momentum.

It is also clear that even for the Holstein model, all expressions \( \Sigma_{\text{MA}}^{(n)} \) with \( n \geq 2 \) will have momentum dependence, since Holstein self-energy diagrams of second order are momentum dependent. We demonstrate below that this is indeed the case.

Finally, the MA approximation becomes exact in the limit \( g \to 0 \) and \( t \to 0 \). The first limit is trivial, since \( \Sigma \to 0 \). The second is due to the fact that if \( t = 0 \) then free propagators \( G_0(k, \omega) \) are in fact independent of \( k \), and thus the momentum averages become irrelevant. Clearly, the same must hold true for all higher level MA\(^{(n)} \) approximations. Indeed, one can verify directly that if \( \bar{g}_0(\omega) = (\omega + i\eta)^{-1} \) (corresponding to \( t = 0 \)), then the expressions for \( \Sigma_{\text{MA}}^{(0)} \) and \( \Sigma_{\text{MA}}^{(1)} \) in Eqs. (2.25) and (3.7), respectively, are equal. The same is true for the MA\(^{(2)} \) results we present below, and all higher orders MA\(^{(n)} \).

3.1.2 MA\(^{(2)} \) Level

In this case, the equations to be solved are:

\[
\Sigma_{\text{MA}}^{(2)}(k, \omega) = \frac{1}{N} \sum_{q_1} f_1^{(2)}(q_1)
\]

(3.11)

where

\[
f_1^{(2)}(q_1) = G_0(k - q_1, \omega - \Omega) \left[ g^2 + \frac{1}{N} \sum_{q_2} f_2^{(2)}(q_1, q_2) \right],
\]

(3.12)

\[
f_2^{(2)}(q_1, q_2) = G_0(k - q_1 - q_2, \omega - 2\Omega)
\times \left( g^2 \left[ f_1^{(2)}(q_1) + f_1^{(2)}(q_2) \right] + \frac{1}{N} \sum_{q_3} f_3^{(2)}(q_1, q_2, q_3) \right),
\]

(3.13)

and for all \( n \geq 3 \),

\[
f_n^{(2)}(q_1, \ldots, q_n) = \bar{g}_0(\omega - n\Omega)
\times \left[ g^2 \sum_{i=1}^{n} f_{n-1}^{(2)}(\ldots, q_{i-1}, q_{i+1}, \ldots) + \frac{1}{N} \sum_{q_{n+1}} f_n^{(2)}(q_1, \ldots, q_{n+1}) \right].
\]

(3.14)
dependence on \( k, \omega \) is again implicitly assumed.

This can be reduced to a closed system of equations in terms of only \( f^{(2)}_1(q_1) \) and \( f^{(2)}_2(q_1, q_2) \), after solving for \( (1/N) \sum_{q_3} f^{(2)}_3(q_1, q_2, q_3) \) from Eq. (3.12). The details are provided in Appendix C.

We use the shorthand notation:

\[
A_1 \equiv A_1(\omega - 2\Omega), \quad A_2 \equiv A_2(\omega - \Omega), \quad A_3 \equiv A_3(\omega).
\] (3.15)

where the continued fractions are defined in Eq. (2.26). We also define various momentum averages (dependence on \( k, \omega \) is again implicit):

\[
F_1 = \frac{1}{N} \sum_{q_1} f^{(2)}_1(q_1),
\] (3.16)

\[
F_2 = \frac{1}{N^2} \sum_{q_1, q_2} f^{(2)}_2(q_1, q_2) = \frac{2g^2 \tilde{g}_0(\tilde{\omega}) F_1}{1 - g^2 \tilde{g}_0(\tilde{\omega})(A_3 - A_1)},
\] (3.17)

and

\[
\delta \tilde{f}_2(q_1) = \frac{1}{N} \sum_{q_2} f^{(2)}_2(q_1, q_2) - F_2.
\] (3.18)

The link between \( F_1 \) and \( F_2 \) is proved in Eq. (C.15). In terms of these, the closed system of equations to be solved becomes (see Appendix C for more details):

\[
f^{(2)}_1(q_1) = G_0(k - q_1, \omega - \Omega) \left[ g^2 + \delta \tilde{f}_2(q_1) + F_2 \right]
\] (3.19)

and

\[
\delta \tilde{f}_2(q_1) = g^2 \tilde{g}_0(\tilde{\omega}) \left[ f^{(2)}_1(q_1) + (A_2 - A_1) \delta \tilde{f}_2(q_1) - 2F_1 \right]
\] 
\[+ \frac{g^2}{N} \sum_{q_2} G_0(k - q_1 - q_2, \tilde{\omega}) \left[ f^{(2)}_1(q_2) + (A_2 - A_1) \delta \tilde{f}_2(q_2) \right].
\] (3.20)

These equations can be solved in a variety of ways. We present here the most efficient solution that we have found, and then comment briefly on other possible solutions. First, given the form of these equations, it is advantageous to introduce the new unknown

\[
x_q = f^{(2)}_1(q) + (A_2 - A_1) \delta \tilde{f}_2(q).
\] (3.21)

Consider its Fourier transform at various lattice sites \( R_i \), namely \( x(i) = (1/N) \sum_q e^{iq \cdot R_i} x_q \). First, observe that \( x(0) = (1/N) \sum_q x_q = F_1 = \Sigma_{\text{MA}^{(2)}} \), by definition.
As shown in Appendix C, the set of two closed equations above can be rewritten as an inhomogeneous equation involving $x(i)$ (at all lattice sites):

$$
\sum_j M_{ij}(k, \omega)x(i) = e^{ik \cdot R_i}g^2 G_0(-i, \tilde{\omega})
$$

where

$$
\tilde{\omega} = \omega - \Omega - \frac{g^2 \bar{g}_0(\tilde{\omega})}{1 - g^2 \bar{g}_0(\tilde{\omega})(A_2 - A_1)}.
$$

$G_0(i, \omega) = (1/N) \sum_k e^{ik \cdot R_i}G_0(k, \omega)$, and the expression of the matrix elements $M_{ij}(k, \omega)$ is given in Eqs. (C.24)-(C.26). They have simple expressions, involving only (the same) three continued fractions $A_1, A_2, A_3$ as well as various $G_0(i, \omega)$ values, therefore they can be calculated easily.

Because at low energies the free propagators in real space decay exponentially, one expects that $x(i)$ decreases fast with increasing distance $R_i$. Alternatively, consider, for instance, the $f^{(2)}_1(q)$ contribution to $x_q$. When Fourier transformed, the initial state $c^\dagger_{k-q}b^\dagger_q|0\rangle$ goes into $c^\dagger_{j}b^\dagger_{j+i}|0\rangle$, i.e. the phonon is further and further apart from the electron. Similar interpretation can be given for the second contribution to $x(i)$. One expects the amplitudes for such processes to decay with $i$.

As a result, we can truncate the system of equations (3.22), assuming that $x(i) = 0$ for all $R_i$ larger than a cutoff. We vary this cutoff to ensure that convergence has indeed been achieved. In this formulation, we find that convergence is reached extremely fast, typically for a cutoff distance of order $5a$ (see results section). In other words, to obtain $\Sigma_{MA}^{(2)}(k, \omega) = x(0)$, in 1D we typically have to solve a system of 11 or so inhomogeneous equations, which can be done very efficiently. Higher dimensions imply larger systems, but overall the numerical task is still trivial and results can be obtained very fast and with little computational resources.

It is important to emphasize that such low cutoffs are not inherent in the problem. In fact, one could also solve these equations, for instance, by eliminating $f^{(2)}_1(q)$ to get an equation only in terms of $\delta f_2(q_1)$ and $F_2 \sim \Sigma$. If one Fourier transforms this, it turns out that cutoffs as large as 100$a$ are needed before convergence is achieved, and this is especially so for the polaron+one-phonon continuum (the bound polaron states converge quickly). This is not surprising, since states in the polaron+one-phonon continuum do have a free phonon, i.e. one that could be infinitely far from the polaron. In reality, we expect that if we have a finite but large enough system, all quantities will eventually converge to their bulk values. In particular, here this suggests
3.1. Higher Level MA

that one needs to allow the free phonon to be hundreds of sites away from
the polaron before convergence for the continuum is achieved.

The much faster convergence for the formulation of Eq. (3.22) is due
in part to the particular choice of variable \( x_q \). Even more important is
the infinite summation of diagrams performed when the new frequency \( \tilde{\omega} \)
appears (see Appendix C). Without this, the convergence for continuum
energies remains slow and large cutoffs are needed. Examples are discussed
in the results section.

Of course, one could also attempt to solve these equations directly in
the \( k \)-space. Without having tried it, we believe this to be an inefficient
approach. The goal is to find \( F_1 = \Sigma \), \textit{i.e.} an average over the Brillouin
zone. Within MA\(^{(1)}\), \( f_1^{(1)}(q) \sim G_0(k - q, \tilde{\omega}) \) (it is a constant in MA).
Presuming that \( f_1^{(2)}(q) \) is not too different, it is clear that these functions are
of comparable size everywhere in the Brillouin zone, and therefore one should
sample many points in the Brillouin zone to obtain an accurate average.

Finally, going back to Eq. (3.22), we would like to point out that if we
set the cutoff at zero distance, \textit{i.e.} use \( M_{00}(0) = g^2 \tilde{g}_0(\tilde{\omega}) \), we obtain an
analytical, momentum-independent approximation to the true \( \Sigma_{MA^{(2)}}(k, \omega) \):

\[
\tilde{\Sigma}_{MA^{(2)}}(\omega) = \frac{g^2 \tilde{g}_0(\tilde{\omega})}{1 - g^2 \tilde{g}_0(\tilde{\omega}) \tilde{g}_0(\tilde{\omega}) \left( \frac{2}{a_{31}(\omega)} - \frac{1}{a_{21}(\omega)} \right)}, \tag{3.24}
\]

where

\[
a_{ij}(\omega) = 1 - g^2 \tilde{g}_0(\tilde{\omega})(A_i - A_j). \tag{3.25}
\]

One can think of this as the variant of MA\(^{(2)}\) that keeps some of the free propagators of energy \( \omega - 2\Omega \) exactly (typically those appearing in non-crossed diagrams), but averages over those that give momentum dependence to the self-energy. The self-energy \( \tilde{\Sigma}_{MA^{(2)}}(\omega) \) is more accurate than \( \Sigma_{MA^{(1)}}(\omega) \) but
less accurate than \( \Sigma_{MA^{(2)}}(k, \omega) \). Such “zero-cutoff” analytical approximations can be obtained for higher levels of MA\(^{(n)}\) quite easily. The full MA\(^{(n)}\)
for \( n > 2 \) can also be done. The reduction to a closed system of \( n \) coupled
equations is always straightforward. Its solution, however, becomes more
computationally involved as \( n \) increases, and leads to gradually smaller
improvements in the accuracy.
3.2 Results

A detailed comparison of the predictions of the MA approximation vs. numerical simulations is given in Section 2.3. Instead of another comprehensive investigation, here we will focus on several properties where the higher level MA\(^{(n)}\) approximations show a significant improvement over MA results. The way of extracting quantities of interest from the Green’s function, e.g. ground state (GS) energies, quasiparticle (qp) weights, effective masses, average number of phonons in the polaron cloud, etc., are described in detail in Chapter 2.

3.2.1 Ground State Properties

The ground state energies predicted by MA are quite accurate for a large range of parameters. The accuracy is known to worsen as \(\Omega/t \to 0\), however even for \(\Omega/t = 0.1\) the MA energies have less than 5\% relative error. On the other hand, the GS qp weight for this low \(\Omega/t\) ratio is quite wrong for intermediary couplings \(\lambda \sim 1\), although it does become asymptotically exact, as expected. The comparison with Quantum Monte Carlo (QMC) results is shown in Fig. 3.1, where we also show the MA\(^{(1)}\) and MA\(^{(2)}\) predictions.

The accuracy is improved significantly for the higher MA levels. Improvements are observed for all other sets of parameters (not shown) analyzed in Section 2.3, however for higher \(\Omega/t\) ratios MA is much more accurate to begin with, so the supplementary improvements due to MA\(^{(1)}\) and MA\(^{(2)}\) are comparatively smaller. As discussed, this systematic improvement is expected since all self-energy diagrams become more and more accurate. This is reflected in the sum rules for spectral weight, which are also systematically improved (the link between diagrams and sum rules is discussed at length in Sections 2.2.1 and 2.2.3). While MA satisfies the first 6 sum rules exactly, MA\(^{(1)}\) satisfies the first 8 sum rules exactly, MA\(^{(2)}\) satisfies the first 10 sum rules exactly, etc. Of course, the accuracy of all higher order sum rules is also systematically improved by the use of more accurate expressions for the diagrams.

In Section 2.2.2 it was shown that MA\(^{(0)}\) is equivalent to a variational approach within a Hilbert space where phonons are only allowed on a single site [see Eq. (2.37)]. From this variational point of view, MA\(^{(1)}\) almost corresponds to using an enlarged variational Hilbert space; after Fourier transforming to real space, it is straightforward to verify that (see Appendix
3.2. Results

![Graph showing ground state energy and quasiparticle weight as a function of effective coupling $\lambda$.](image)

Figure 3.1: (a) Ground state energy and (b) Ground state quasiparticle weight, as a function of the effective coupling $\lambda$, for $d = 1, t = 1, \Omega/t = 0.1$. The QMC results are from Ref. [26].

\[ \delta f_n(q_1) \sim \sum_{i \neq j} e^{i(k-q_1) \cdot R_i + i q_1 \cdot R_j} c_i (b_i^\dagger)^{n-1} b_j^\dagger |0\rangle, \]  

(3.26)

\[ i.e. \] one phonon is now allowed to be away from the main polaron cloud. The equivalence is not exact, because the resulting MA$^{(1)}$ equations do not contain a term that would appear in a full variational treatment, so more correctly one could view this as also being accompanied by a change in the effective Hamiltonian. Similarly, MA$^{(2)}$ allows for up to two phonons to be located anywhere away from the main phonon cloud, again with a slight change of the effective Hamiltonian when acting on these isolated phonons.

This systematic increase in the size of the variational Hilbert space is another way to explain the gradual improvement of the GS energy. Also, it is now clear that the polaron + one-phonon continuum should appear at level MA$^{(1)}$ (see below). As a result, the GS qp weight no longer needs to account for it and it decreases, improving the agreement with QMC results as shown in Fig. 3.1(b). We will return to this issue when we investigate spectral weights. For the time being, we note that in the adiabatic limit.
3.2. Results

$\Omega/t \to 0$, many phonons can be created in the system at low energetic cost. In the intermediary region $\lambda \sim 1$ where the polaron cloud is still relatively large, one expects many of these phonons to be relatively far from the polaron and therefore a high order $n$ would be required in order to accurately describe them within this approach. As a result, it is expected that the strongly adiabatic regime will not be quantitatively well described for $\lambda \sim 1$ by the low-level MA approximations, even though the qualitative behavior is correctly captured. Of course, this limit can be investigated by other means, such as in Ref. [38] and references therein. However, for most of the parameter space, i.e. any $\Omega/(dt) > 0.1$ or so and any coupling $\lambda$, the MA set of approximations give very easy to evaluate yet remarkably accurate results.

3.2.2 Polaron Band

We can also track how the lowest eigenstate of momentum $k$, and its various properties, change with various parameters. Results are shown in Figs. 3.2 and 3.3 for the energy $E_k$, $qp$ weight $Z_k$ and average number of phonons $N_{ph}(k)$ for 1D and two couplings, $\lambda = 0.25$ and $\lambda = 1.00$. We found similar improvements in 2D cases. Clearly, MA$^{(2)}$ leads to an obvious improvement, even though for this value of $\Omega/t = 0.5$, MA itself is quite accurate already.

It should be noted that $\lambda \sim 1$ is where the MA accuracy is generally expected to be at its worst. In particular, for the weak-coupling value $\lambda = 0.25$, we see that MA overestimates the distance to the continuum, i.e. $E_\pi - E_0 > \Omega$. This should be $\Omega$, but it is larger for MA because the polaron+one-phonon continuum is not predicted at the correct energy. For the intermediate coupling $\lambda = 1$ there is a second bound state between the one shown and the continuum, therefore the bandwidth is much less than $\Omega$. We see that at the MA$^{(1)}$ and MA$^{(2)}$ levels this problem is indeed fixed, and the polaron dispersion width (at weak couplings) is $\Omega$. All other quantities are also clearly more accurate.

3.2.3 Higher Energy States

In order to understand the effects on higher-energy states, we study the spectral weight $A(k, \omega) = -(1/\pi) \text{Im} G(k, \omega)$. As is well known, this is finite only at energies $\omega$ where eigenstates of momentum $k$ exist. For discrete (bound) states the spectral weight is a Lorentzian of width $\eta$ [see Eq. (1.15)] and area equal to the $qp$ weight. In a continuum, the lifetime is determined by $\text{Im}\Sigma(k, \omega)$ and is independent of $\eta$, provided that $\eta$ is chosen small enough.
3.2. Results

Figure 3.2: (a) and (b) Polaron dispersion $E_k$; (c) and (d) $qp$ weight $Z_k$ vs. $k$, in $d = 1$, for $\Omega = 0.5t$ and $\lambda = 0.25$ in (a), (c), respectively $\lambda = 1.00$ in (b),(d). The QMC results are from Ref. [26].
3.2. Results

Figure 3.3: (a) and (b) Average number of phonons $N_{\text{ph}}(k)$ vs. $k$, in $d = 1$, for $\Omega = 0.5t$ and $\lambda = 0.25$ in (a) and $\lambda = 1.00$ in (b). The QMC results are from Ref. [26].
3.2. Results

Figure 3.4: (a) Spectral weight $A(k = 0, \omega)$ vs. $\omega$ in 1D for $t = 1, \Omega = 0.5, \lambda = 0.6, \eta = 0.01$, in MA, MA$^{(1)}$ and MA$^{(2)}$ (curves are shifted for clarity); (b) Polaron+one-phonon continuum convergence with cutoff within MA$^{(2)}$. Results for cutoffs of 0, 1, 3, 5 and 10 are shown (the last two are almost identical). For comparison, the MA$^{(1)}$ continuum is also shown (black full line); (c) same as in (b), but for an inefficient computation scheme.

In Fig. 3.4(a) we show results for the 1D spectral weight $A(k = 0, \omega)$ vs. $\omega$ for a relatively weak coupling $\lambda = 0.6$. The MA spectral weight shows two discrete states at low energies, and a continuum starting for $\omega > -1.5t$. Within MA$^{(1)}$, the second peak spreads into a continuum whose lower edge is at roughly $\Omega$ above the energy of the GS peak. In fact, since $\tilde{g}_0(\omega)$ acquires an imaginary part when $-2dt \leq \omega \leq 2dt$, from Eq. (3.7) it follows that the MA$^{(1)}$ continuum appears when

$$\tilde{\omega} \geq -2t \rightarrow \omega > E_0^{(0)} + \Omega,$$

(3.27)

where $E_0^{(0)}$ is the MA prediction for the GS energy. Since the MA$^{(1)}$ GS energy $E_0^{(1)} < E_0^{(0)}$, it follows that the gap is in fact slightly larger than $\Omega$. In MA$^{(2)}$ the weight is redistributed within this continuum, and its lower band-edge is at $\Omega$ above the ground state energy, within numerical precision.

The convergence of MA$^{(2)}$ with the cutoff value used to truncate Eqs. (3.22) is shown in Fig. 3.4(b). For a cutoff value of 0 we obtain the momentum-independent self-energy of Eq. (3.24), which gives a continuum with a shape rather similar to that predicted by MA$^{(1)}$. As the cutoff
value is increased, weight shifts towards the lower band-edge. Convergence is reached very quickly, with little difference visible between results corresponding to a cutoff of 5 or 10 [these imply solving an inhomogeneous system of 11, respectively 21 equations in Eqs. (3.22)]. Other solutions of the coupled equations (briefly discussed in the previous section) converge much more slowly. In Fig. 3.4(c) we show results for 3 cutoffs for such an alternative scheme. Even for a cutoff as large as 100, one can still see small oscillations, very reminiscent of finite-size effects. The finding of the efficient solution for the \( \text{MA}^{(2)} \) self-energy is thus quite important.

Similar behavior is observed for higher couplings, as seen in Figs. 3.5(a) and 3.6(a) for intermediate, respectively strong couplings \( \lambda = 1.2 \) and 1.8. In both cases there are now two bound, discrete states below the continuum starting at \( E_0 + \Omega \), as expected [53]. The weight of the polaron+one-phonon continuum decreases very fast, so that for \( \lambda = 1.8 \) it is barely visible just above the second peak. Fig. 3.6(b) shows it clearly, on a logarithmic scale. Interestingly, it is not only the height of this feature that is much smaller as \( \lambda \) increases, but its width as well. In fact, scaling vs. \( \eta \) in Fig. 3.6(b) shows that this is more like a Lorentzian, i.e. a single bound state, and not a finite-width continuum as was the case for lower couplings [see Fig. 3.5(b)]. This
3.2. Results

Figure 3.6: (a) Spectral weight $A(0, \omega)$ vs. $\omega$ in 1D for $t = 1, \Omega = 0.5, \lambda = 1.8, \eta = 0.01$, in MA, MA$^{(1)}$ and MA$^{(2)}$ (curves shifted for clarity); (b) $\ln A(0, \omega)$ vs. $\omega$ for MA$^{(2)}$ and $\eta = 10^{-2}, 10^{-3}, 10^{-4}$. Other parameters are as in (a).

is in fact reasonable, since at such large couplings the lowest energy polaron state is basically dispersionless (the effective mass is already considerable and the polaron is well into the small-polaron regime). Since the width of the continuum is due to the polaron dispersion (the phonon being dispersionless) it is reasonable that as the polaron bandwidth decreases exponentially with increasing coupling, so does the width of the polaron+one-phonon continuum.

The higher energy features are also quite interesting. For the intermediate coupling $\lambda = 1.2$, at some distance above the polaron+one-phonon continuum one can see the feature evolved from what was the third discrete state in the MA approximation. For a large $\eta$ value this looks like a continuum, however scaling with $\eta$ reveals a discrete state just below another continuum. In fact, the spectrum is broken up into discrete states and continua separated by gaps where no states are present. Of course, the detailed shape of the spectral weight above the third bound state is likely to change as one goes to MA$^{(3)}$ and higher orders, however we doubt that these gaps should all close and a single continuum should form above $E_0 + \Omega$, as is the case at very weak couplings (see below). Instead, to us these results suggest that most weight is inside bound states which are reminiscent of the
3.2. Results

Figure 3.7: (left) MA spectral weight $A(0, \omega)$ vs. $\omega$ in 1D for $t = 1, \Omega = 0.5, \eta = 0.01$ and $\lambda$ varying from 0 to 2; (right) Same for MA$^{(2)}$. Curves corresponding to $\lambda = 0.3, 0.6, 1.2$ and 1.8 are highlighted.

Lang-Firsov “comb” of discrete states separated by a frequency $\Omega$. Here the distance between discrete states is generally less than $\Omega$ and there are narrow continua in between them, which evolve from lower-energy bound-states + one or more phonons. This is nicely illustrated by the results in Fig. 3.6(b), which show 2 different features between the third and the fourth bound states. The lower-energy one is a continuum that starts roughly at $\Omega$ above the second bound state, and has a finite width. Since the second bound state still has some finite bandwidth at this coupling (see below), it seems reasonable to interpret this feature as being the polaron in the second-bound state plus one phonon somewhere far from it. The higher energy feature is much narrower, but close inspection reveals that it also is a continuum, with the lower edge starting at $E_0 + 2\Omega$, so its origin is obvious. The continua between higher bound states become wider, in agreement with expectations if one assumes that indeed they result from adding a distant phonon to a polaron in a higher bound state, which has a larger bandwidth. There is also overlap between different types of states, which also leads to increased bandwidth.

This structure of the spectral weight explains why the MA (which predicts only bound states at low energies, for medium to strong coupling) still obeys sum rules with such good accuracy. Most of the weight is indeed in
3.2. Results

the bound states, not in the narrow continua that appear in between them. These results also show clearly how the convergence towards the Lang-Firsov limit \( g \gg t \) is achieved: the width and weight of the continua shrinks to zero, and one is left only with the equally-spaced discrete states. This is illustrated in Fig. 3.7, where \( k = 0 \) MA spectral weights are contrasted with MA\(^{(2)}\) spectral weights for different \( \lambda \) values. The largest difference is observed at small couplings, where MA predicts the (wrong) continuum pinned at \(-2dt + \Omega = -1.5\) for these values, whereas MA\(^{(2)}\) clearly shows a continuum starting at \( \Omega \) above the ground state for as long as it is still visible. For \( \lambda > 0.7 \) or so, a second bound state splits from this continuum and comes quite close to the GS peak before moving away so that it asymptotically goes to \( E_0 + \Omega \). The existence of second bound state below the first continuum, for moderate and strong couplings, is well known for Holstein polaron and has also been demonstrated for other polaron models [53, 84]. It further serves to validate these approximations. There are clear similarities between the two plots, with most of the weight concentrated in the bound states that are roughly \( \Omega \) apart, but MA overestimates their weights in order to compensate for the missing narrow continua that appear in between these discrete states. As stated, the precise shape and weight of the higher continua is very likely to change as one goes to a higher level MA approximation, however we expect the general picture to remain the same qualitatively.

In passing, we note that although the appearance of a second bound state below the first continuum is commonly found in the Holstein model at moderate to large couplings, the appearance of more than two bound states has not been reported in the literature. In Fig. 3.8 we plot the spectral weight using MA\(^{(2)}\) in 3D and show that a third bound state appears below the first continuum. This bound state is only visible over a very narrow range of el-ph couplings around \( \lambda = 1.05 \), and it appears at both the MA\(^{(1)}\) and MA\(^{(2)}\) levels of approximation. Although there is no physical process that prevents the formation of three or more bound states below the continuum (that we are aware of), this feature is still interesting in its own right as it has not been reported before. This result is currently under investigation as part of a 3D QMC study of the Holstein model [85].

We are unable to check quantitatively the accuracy of the higher energy spectral weight against detailed numerical predictions, beyond the sort of comparisons shown for the polaron dispersion in Figs. 3.2 and 3.3. The reason is that most of the numerical work is focused on computing low-energy properties. The much fewer high-energy results, such as those based on a variational treatment in Ref. [62] and a novel QMC / exact diago-
3.2. Results

Figure 3.8: Spectral weight $A(0, \omega)$ vs. $\omega$ in 3D for $t = 1$, $\Omega = 0.5$, $\lambda = 1.05$, and $\eta = 0.01, 0.001$ using MA$^{(2)}$, shown on a logarithmic scale. The vertical lines indicate the ground state and the polaron+one-phonon continuum. A third bound state is found to appear below the continuum for intermediate couplings $\lambda \sim 1$ in 3D.

nalization approach in Ref. [49] use a rather large $\eta$ and are already in reasonable agreement with MA, as discussed in the previous chapter. Cluster perturbation theory (CPT) results such as shown in Ref. [66], while also an approximation, are in good qualitative and quantitative agreement with ours. It would be very interesting to be able to compare our results against detailed high-accuracy, high-energy numerical predictions.

Finally, we contrast the difference between MA and MA$^{(2)}$ spectral weights for different momenta and energies. Typical results are shown in Figs. 3.9 and 3.10, for weak, intermediate and strong couplings $\lambda = 0.3, 0.6, 1.2$ and 1.8, respectively. For the weak coupling, as already discussed, the most obvious difference is that the polaron bandwidth is decreased to its correct value of $\Omega$ in MA$^{(2)}$. The qp weights and all other features are very similar. Based on the MA$^{(2)}$ results, it is now clear that the strong resonance seen in the electron+phonon continuum occurs at an energy of $2\Omega$ above the ground state energy. One expects that here is where the second bound state will arise from. For $\lambda = 0.6$, Figs. 3.9(c) and (d)
3.2. Results

show a bigger contrast. Here, MA already predicts a second bound state that has evolved in between the polaron band and the higher-energy continuum. In contrast, MA\(^{(2)}\) shows that there is no second-bound state yet, however the electron+one-phonon continuum is split off the higher energy continuum, which also starts to fractionalize at higher energies (roughly multiples of \(\Omega\)). Within MA\(^{(2)}\), a true second bound-state is observed for the higher couplings shown in Fig. 3.10. In contrast to MA, which shows several bound states which disperse as \(k\) increases, MA\(^{(2)}\) also clearly shows continua in between these discrete states. These account for some of the spectral weight that was in the MA peaks. These results again suggest a very fractionalized spectrum at intermediate and strong couplings. Instead of the polaron band, a second-bound state and a rather featureless continuum at higher energies,
3.2. Results

Figure 3.10: $A(k,\omega)$ vs. $k$ and $\omega$ in 1D for $t = 1, \Omega = 0.5, \eta = 0.01$ and $\lambda = 1.2$ in (a), (b) and $\lambda = 1.8$ in (c), (d). Results for MA are shown in (a), (c), while $\text{MA}^{(2)}$ is shown in (b), (d).

we instead find many sets of discrete states interspersed with continua. As $\lambda \to \infty$, the relative weight in these continua decreases and the spectral weights evolve continuously towards the Lang-Firsov set of discrete states with energies $-g^2/\Omega + n\Omega$.

Qualitatively similar results are observed in higher dimensions. Rather ironically, the most time-consuming part in the $\text{MA}^{(2)}$ calculation for higher dimensions was finding the spatial dependence of the non-interacting Green’s functions $G_0(i,\Omega)$ which are needed to generate Eqs. (3.22), and not the solving of the system of equations. The reason is that for nearest-neighbour hopping in higher dimensions, the evaluation of these propagators was done numerically. Of course, one could choose simpler forms of the dispersion

92
3.2. Results

Figure 3.11: (a) $A(0, \omega)$ vs. $\omega$ in 2D for $t = 1, \Omega = 0.5, \lambda = 0.3, \eta = 0.01$, in MA, MA$^{(1)}$ and MA$^{(2)}$ (curves shifted for clarity); (b) $\ln A(0, \omega)$ vs. $\omega$ for MA$^{(2)}$ and $\eta = 10^{-2}, 10^{-3}, 10^{-4}$. Other parameters are as in (a).

$\varepsilon_k$ for which analytical results are possible. However, as we show now, in higher dimensions the improvements in going to MA$^{(1)}$ and MA$^{(2)}$ are quantitatively smaller, because the relative weight in the continua is reduced compared to the 1D case. This is in agreement with our general observation that MA itself becomes more accurate with increased dimensionality.

In Figs. 3.11-3.14 we show the 2D spectral weight $A(k = 0, \omega)$ vs. $\omega$, for effective couplings $\lambda = 0.3, 0.6, 0.95$ and 1.2, both on linear and logarithmic scales. Qualitatively, everything is similar to the behavior seen in the 1D case. Quantitatively, we find that the continuum that appears at $\Omega$ above the ground state is broader, but of lower height. In fact, its height is so small that it is invisible on curves like those in Fig. 3.7, which is the reason why we do not show such curves here. The overall weight in this continuum also decreases much faster with increasing $\lambda$. In 1D, for $\lambda = 1.2$ the first continuum is still clearly visible (see Fig. 3.5). For $\lambda = 1.8$ it becomes harder to see on the linear scale, but it is clearly seen on the logarithmic scale. By contrast, in 2D, for $\lambda = 1.2$ the continuum is no longer visible on the linear scale, and even in the logarithmic scale it only barely starts to be visible for $\eta = 10^{-4}$ (small shoulder marked by arrow). The spectral weight for a Lorentzian contribution $Z/(\omega + i\eta)$ is $Z\eta/\pi(\omega^2 + \eta^2)$, so the
3.2. Results

![Graphs](image)

Figure 3.12: (a) $A(0, \omega)$ vs. $\omega$ in 2D for $t = 1, \Omega = 0.5, \lambda = 0.6, \eta = 0.01$, in MA, MA$^{(1)}$ and MA$^{(2)}$ (curves shifted for clarity); (b) ln $A(0, \omega)$ vs. $\omega$ for MA$^{(2)}$ and $\eta = 10^{-2}, 10^{-3}, 10^{-4}$. Other parameters are as in (a).

![Graphs](image)

Figure 3.13: (a) $A(0, \omega)$ vs. $\omega$ in 2D for $t = 1, \Omega = 0.5, \lambda = 0.95, \eta = 0.01$, in MA, MA$^{(1)}$ and MA$^{(2)}$ (curves shifted for clarity); (b) ln $A(0, \omega)$ vs. $\omega$ for MA$^{(2)}$ and $\eta = 10^{-2}, 10^{-3}, 10^{-4}$. Other parameters are as in (a).
3.2. Results

Figure 3.14: (a) $A(0, \omega)$ vs. $\omega$ in 2D for $t = 1, \Omega = 0.5, \lambda = 1.2, \eta = 0.01$, in MA, MA$^1$ and MA$^2$ (curves shifted for clarity); (b) $\ln A(0, \omega)$ vs. $\omega$ for MA$^2$ and $\eta = 10^{-2}, 10^{-3}, 10^{-4}$. Other parameters are as in (a).
maximum height of the peak, at resonance, is $Z/\pi \eta$. This is visible only if it is larger than the background, which of course depends on how close is the next spectral feature. However, for a very small $Z$, $\eta$ has to be very small before the peak is seen.

As in the 1D case, we also observe the fractionalization of the spectrum for moderate and large effective couplings, with a succession of discrete peaks and continua at higher energies. This was found in 3D as well (see Fig. 3.8). Clearly, the changes in going from MA to MA$^{(1)}$ and MA$^{(2)}$ are quantitatively much smaller in 2D than in 1D, because the continua have so little weight. We expect the trend to continue in going to 3D, meaning that in 3D, the difference between MA$^{(2)}$ and MA should be quantitatively even less. Indeed, all the comparisons of MA results with available 3D numerics, shown in the previous chapter, are already in excellent agreement, even for intermediary couplings $\lambda \sim 1$. As a result, we find it unnecessary to present 3D results, although they can be obtained straightforwardly.

3.3 Conclusion

In summary, in this chapter, we presented a way to systematically improve the MA approximation, by systematically improving the accuracy of self-energy diagrams, in such a way that they can still all be efficiently summed.

This allows us to rather easily fix various known failings of the MA approximation, such as the absence of the polaron+one-phonon continuum at the correct energy, and its momentum-independent self-energy. It also allows us to understand in more detail the effects of the Holstein-type electron-phonon coupling on the polaron spectrum, both at low and high energies. While agreement with exact numerical results is improved, unfortunately there are not many such results for higher energy states, so detailed comparisons are not possible there. However, the hierarchy of MA$^{(n)}$ approximations is clearly providing a simple way towards obtaining quantitatively more and more accurate results for the Green’s function of the Holstein polaron, in any dimension and for any free electron dispersion.

The next direction of obvious interest is to study to what extent this work can be extended to other models, for example models where the el-ph coupling depends on the phonon momentum. For such models there are very few reliable high-energy results available. A simple approximation such as MA could easily investigate the whole parameter space and identify interesting regimes. The generalization of MA to such models, and in particular the breathing-mode Hamiltonian, will be the subject of the next chapter.
Chapter 4

Generalization to Models with Electron-Phonon Coupling Dependent on the Phonon Momentum

In this chapter we generalize this powerful set of approximations to a much broader class of models with electron-phonon coupling that depends on the phonon momentum. We first derive a simple generalization, leading to what we will continue to call the MA\(^{(0)}\), MA\(^{(1)}\), MA\(^{(2)}\), etc. hierarchy. These are very easy to apply to any Hamiltonian of this class, however while still asymptotically exact for both weak and strong coupling, at intermediary couplings the relative errors are of a few percent. In other words, these can be used to get a quick estimate of typical energies and spectra. We then show how these can be significantly improved, using variational considerations. This generates a second hierarchy of approximations which we call MA\(^{(v,0)}\), MA\(^{(v,1)}\), MA\(^{(v,2)}\), etc., with relative errors well below one percent. As a test case to gauge these accuracies, we use the one-dimensional (1D) breathing-mode Hamiltonian, where high accuracy numerical results have recently become available [22].

This chapter is organized as follows. In Section 4.1 we describe the general el-ph coupling model that we will investigate and in Section 4.2 we derive the simple straightforward generalizations MA\(^{(0)}\), MA\(^{(1)}\), etc. for this model. In Section 4.3 we show how to obtain the more accurate MA\(^{(v,0)}\) and MA\(^{(v,1)}\) approximations based on variational ideas. Here we use the 1D breathing-mode Hamiltonian as an explicit example. In Section 4.4 we present our results and compare them to the available numerical data where possible. Finally, Section 4.5 contains our summary and conclusions for this chapter.
4.1 The Model

The general el-ph coupling model that we consider has a single dispersionless phonon branch, and takes the following form in momentum space [see Eq. (1.10)]:

\[
\mathcal{H} = \sum_k \varepsilon_k c_k^\dagger c_k + \Omega \sum_q b_q^\dagger b_q + \sum_{k,q} \frac{g_q}{\sqrt{N}} c_{k-q}^\dagger c_k \left(b_q^\dagger + b_q\right). \tag{4.1}
\]

The notation is the same as for the Holstein model, and the key difference is that here we allow for a more general momentum-dependent coupling \(g_q\). As discussed in Section 1.5, this general Hamiltonian covers complicated el-ph couplings such as those found in the Fröhlich and breathing-mode models, and it reduces to the Holstein model when \(g_q\) is simply a constant.

In this chapter we will primarily use the 1D breathing-mode Hamiltonian as an example. This model has the form of Eq. (4.1) with

\[
g_q = -2i \sin \frac{qa}{2}, \tag{4.2}
\]

however, we stress that the methods we present are applicable to any Hamiltonian like Eq. (4.1) in any dimension. As discussed in Section 1.5, the reason for choosing this model is two-fold. First, in its full 2D form it describes the lattice vibrations in a CuO\(_2\)-like plane, where the motion of the O ions living on the bonds connecting the Cu sites is the most important vibrational degree of freedom, making it possibly relevant for the study of high-\(T_c\) superconductors. The second reason is that exact diagonalization (ED) results have recently become available for its 1D analog [22]. These results serve as an excellent gauge of the accuracy of the generalized MA approximations, and their availability is very fortunate because although there are many numerical results for the Holstein model, it is only due to recent advancements in computational power that more complicated el-ph coupling models, such as the breathing-mode Hamiltonian, can be investigated numerically.

4.2 Calculating the Green’s Function

4.2.1 Exact Solution

In Section 2.1.1 we described in detail how to generate these equations for the Holstein model. The generalization to momentum-dependent coupling models is straightforward, again through the repeated use of Dyson’s identity:

\[
\hat{G}(\omega) = \hat{G}_0(\omega) + \hat{G}(\omega) \hat{V} \hat{G}_0(\omega), \tag{4.3}
\]
where \( \hat{G} = [\omega - \hat{H} + i\eta]^{-1} \), \( \hat{G}_0 = [\omega - \hat{H}_0 + i\eta]^{-1} \), \( \hat{H} = \hat{H}_0 + \hat{V} \), and \( \hat{V} \) is the el-ph interaction. Recalling the definition of the generalized Green’s functions

\[
F_n(k, q_1, \ldots, q_n, \omega) = \langle 0 | c_k \hat{G}(\omega) e_{k-q_T}^\dagger b_{q_1}^\dagger \ldots b_{q_n}^\dagger | 0 \rangle, \tag{4.4}
\]

where \( F_0(k, \omega) = G(k, \omega) \), we find that

\[
G(k, \omega) = G_0(k, \omega) \left[ 1 + \frac{1}{\sqrt{N}} \sum_{q_1} g_{q_1} F_1(k, q_1, \omega) \right], \tag{4.5}
\]

and for \( n \geq 1 \),

\[
F_n(k, q_1, \ldots, q_n, \omega) = \frac{1}{\sqrt{N}} G_0(k - q_T, \omega - n\Omega) \times \left[ \sum_{i=1}^{n} g_{-q_i} F_{n-1}(k, \ldots, q_{i-1}, q_{i+1}, \ldots, \omega) \right. \\
\left. + \sum_{q_{n+1}} g_{q_{n+1}} F_{n+1}(k, q_1, \ldots, q_{n+1}, \omega) \right]. \tag{4.6}
\]

The total momentum carried by the phonons is denoted by \( q_T = \sum_{i=1}^{n} q_i \) and \( G_0(k, \omega) = (\omega - \varepsilon_k + i\eta)^{-1} \) is the free electron Green’s function. Observing from Eqs. (4.6) that all of these generalized Green’s functions \( F_1, F_2, \ldots \) must be proportional to \( G(k, \omega) \), we can again recast our equations into a more convenient form by defining

\[
f_n(q_1, \ldots, q_n) = \frac{N^{n/2} g_{q_1} \cdots g_{q_n} F_n(q_1, \ldots, q_n)}{G(k, \omega)}, \tag{4.7}
\]

where we have also introduced the shorthand notation \( f_n(k, q_1, \ldots, q_n, \omega) \equiv f_n(q_1, \ldots, q_n) \) (i.e., the \( k \) and \( \omega \) dependence of these functions is implicitly assumed from now on). In this notation, Eqs. (4.5) and (4.6) become

\[
G(k, \omega) = G_0(k, \omega) \left[ 1 + \frac{1}{N} \sum_{q_1} f_1(q_1) G(k, \omega) \right], \tag{4.8}
\]

99
4.2. Calculating the Green’s Function

\[ f_n(q_1, \ldots, q_n) = G_0(k - q_T, \omega - n\Omega) \]
\[ \times \left[ \sum_{i=1}^{n} |g_{q_i}|^2 f_{n-1}(\ldots, q_{i-1}, q_{i+1}, \ldots) ight. \]
\[ + \frac{1}{N} \sum_{q_{n+1}} f_{n+1}(q_1, \ldots, q_{n+1}) \right], \]  
(4.9)

with a solution written in the standard form

\[ G(k, \omega) = \frac{1}{\omega - \varepsilon_k - \Sigma(k, \omega) + i\eta}, \]  
(4.10)

where the self-energy is

\[ \Sigma(k, \omega) = \frac{1}{N} \sum_{q_1} f_1(q_1). \]  
(4.11)

It is easy to see that these equations reduce to Eqs. (2.8)-(2.11) when \( g_q \) is replaced by a constant. This self-energy can also again be written in terms of an infinite set of diagrams, as shown in Fig. 2.1.

Of course this system can be solved trivially in the limit of \( g = 0 \), in which case \( G(k, \omega) = G_0(k, \omega) \) directly from Eq. (4.5). Also, in the limit of \( t = 0 \) the free propagators become independent of momentum and the equivalent of the Lang-Firsov result is reproduced as before. However, as was the case for the Holstein model, a closed form solution cannot be obtained, and approximations are therefore needed. We begin by describing the simplest \( \text{MA}^{(0)} \) version of a generalized MA approximation.

4.2.2 The \( \text{MA}^{(0)} \) Approximation

For the Holstein model, the \( \text{MA}^{(0)} \) approximation amounts to replacing all of the free electron propagators in the diagrammatical expansion of the self-energy by their momentum average over the Brillouin zone:

\[ \bar{g}_0(\omega) = \frac{1}{N} \sum_q G_0(q, \omega), \]  
(4.12)

which is equivalent to replacing all \( G_0(k - q_T, \omega - n\Omega) \) by \( \bar{g}_0(\omega - n\Omega) \) in Eq. (4.9). The procedure is essentially the same for momentum-dependent el-ph
4.2. Calculating the Green’s Function

couplings. We note that the first term from the \( n = 1 \) case of Eq. (4.9) does not actually require any approximation (this is trivially true in the case in the Holstein model) because \( f_0 \equiv 1 \) by definition, and its coefficient can be written explicitly as

\[
\bar{g}_0(k, \omega) = \frac{1}{N} \sum_q |g_q|^2 G_0(k - q, \omega).
\] (4.13)

By making the substitution \( G_0(q, \omega - n\Omega) \rightarrow \bar{g}_0(\omega - n\Omega) \) everywhere else, we can rewrite Eqs. (4.9) and (4.11) in terms of the momentum averaged functions

\[
\mathcal{F}_n = \frac{1}{N^n} \sum_{q_1, \ldots, q_n} f_n(q_1, \ldots, q_n).
\] (4.14)

This leads to simple recurrence relations linking each \( \mathcal{F}_n \) to \( \mathcal{F}_{n-1} \) and \( \mathcal{F}_{n+1} \), and these can be solved in terms of continued fractions like before. We find

\[
\Sigma_{MA}^{(n)}(k, \omega) = \frac{\bar{g}_0(k, \omega - \Omega)}{1 - \bar{g}^2 \bar{g}_0(\omega - n\Omega) A_n(\omega)},
\] (4.15)

where [see Eq. (2.26)]

\[
A_n(\omega) = \frac{n \bar{g}_0(\omega - n\Omega)}{1 - \bar{g}^2 \bar{g}_0(\omega - n\Omega) A_{n+1}(\omega)},
\] (4.16)

and we have defined the momentum averaged el-ph coupling:

\[
\bar{g}^2 = \frac{1}{N} \sum_q |g_q|^2.
\] (4.17)

It is trivial to check that this reduces to the Holstein result of Eq. 2.25 when \( g_q \) is independent of momentum.

The meaning of the approximation in this more general case is the same as before (see Section 2.2), and we don’t repeat the justification for the MA approximation here. The approximation can again be systematically improved. In Chapter 3 we derived explicitly the Holstein MA self-energy for both \( n = 1 \) and \( n = 2 \). As shown and explained there, the \( n = 1 \) order already ensures the key improvement of properly predicting the polaron+one phonon continuum, which is usually absent at the \( n = 0 \) level. The reason we went to \( n = 2 \) for Holstein is that, given the simplicity of that model, only at the MA\(^{(2)}\) level did we find a momentum-dependent self-energy. However, for a model with a momentum-dependent coupling even the MA\(^{(0)}\) level gives explicit \( k \) dependence in the self-energy [see Eq. (4.15)], therefore in the next subsection we only consider the MA\(^{(1)}\) generalization.
4.2. Calculating the Green’s Function

4.2.3 The MA(1) Approximation

The main drawback of the MA(0) level of approximation is that it predicts an incorrect location for the electron+phonon continuum that must start at \( E_0 + \Omega \). This problem is always “cured” at the MA(1) level because in real space MA(1) includes states with a phonon cloud near the electron and a single phonon arbitrarily far away, in other words precisely the type of states that give rise to the polaron+phonon continuum. Mathematically, MA(1) amounts to keeping all propagators with the argument \( \omega - \Omega \) in Eq. (4.9) exactly. Such terms only appear in the \( n = 1 \) equation:

\[
f_1^{(1)}(q_1) = G_0(k - q_1, \omega - \Omega) \left[ |g_{q_1}|^2 + \frac{1}{N} \sum_{q_2} f_2^{(1)}(q_1, q_2) \right],
\]

where we have distinguished the approximated \( f_n \) terms with the superscript (1) to indicate the MA(1) level of approximation. For the remaining equations \( (n \geq 2) \) we proceed as before, replacing \( G_0(k - q_T, \omega - n\Omega) \) with \( g_0(\omega - n\Omega) \) everywhere in Eq. (4.9):

\[
f_n^{(1)}(q_1, \ldots, q_n) = g_0(\omega - n\Omega)
\times \left[ \sum_{i=1}^{n} |g_{q_1}|^2 f_{n-1}^{(1)}(\ldots, q_{i-1}, q_{i+1}, \ldots)
+ \frac{1}{N} \sum_{q_{n+1}} f_{n+1}^{(1)}(q_1, \ldots, q_{n+1}) \right].
\]

We wish to solve for \( \Sigma_{\text{MA}^{(1)}}(k, \omega) = (1/N) \sum_{q_1} f_1^{(1)}(q_1) \). The procedure is analogous to that presented in Section 3.1.1. We obtain two sets of coupled recurrence relations, one for fully momentum averaged quantities

\[
F_n = \frac{1}{N^n} \sum_{q_1, \ldots, q_n} f_n^{(1)}(q_1, \ldots, q_n)
\]

that already appeared at MA(0) level, and one for partially averaged quantities

\[
\bar{f}_n(q_1) = \frac{1}{N^{n-1}} \sum_{q_2, \ldots, q_n} f_n^{(1)}(q_1, \ldots, q_n).
\]

Their solution follows the procedure detailed in Section 3.1.1, and we simply state the result:

\[
\Sigma_{\text{MA}^{(1)}}(k, \omega) = \frac{\bar{g}_0(k, \tilde{\omega})}{1 - \bar{g}_0(k, \tilde{\omega}) [A_2(\omega) - A_1(\omega - \Omega)]}.
\]
4.2. Calculating the Green’s Function

where \( \tilde{\omega} = \omega - \Omega - \bar{g}^2 A_1(\omega - \Omega) \). This expression is slightly more complicated than the expression for \( \Sigma_{\text{MA}(\omega)}(\mathbf{k}, \omega) \) because it involves two continued fractions, but it is again trivial to compute numerically.

4.2.4 MA\(^{(0)}\) and MA\(^{(1)}\) Results

To illustrate the accuracy of MA\(^{(0)}\) and MA\(^{(1)}\) we will use the 1D breathing-mode Hamiltonian described by Eqs. (4.1) and (4.2) as an example, comparing our results to those found numerically using ED in Ref. [22].

With our approximations for the Green’s function in hand, the ground-state properties are found by tracking the energy and weight of the lowest pole of the spectral weight. The explicit expressions for \( \bar{g}_0(\omega) \) and \( \bar{g}_0(k, \omega) \) needed to compute the self-energies for the 1D breathing-mode Hamiltonian are given in Appendix D. As is customary, we use the dimensionless effective coupling \( \lambda \) as the ratio between the lattice deformation energy \( -2g^2/\Omega \) and the free electron ground state energy \( -2t \), given by

\[
\lambda = \frac{g^2}{\Omega t}.
\]

(4.23)

Recall that the lattice deformation energy, and therefore \( \lambda \), is two times larger for the breathing-mode polaron than for the Holstein polaron, as shown in Sections 1.4.1 and 1.5.1.

In Fig. 4.1 we plot the ground-state energy and qp weight as a function of the el-ph coupling strength, for a phonon energy \( \Omega/t = 0.5 \). The MA\(^{(0)}\) (red dashed line) and MA\(^{(1)}\) (solid blue line) results both show good agreement with the ED results (black circles). The ground state energies calculated with MA\(^{(0)}\) are within 5% error of the ED results, and the MA\(^{(1)}\) results are better, coming within 3% error of the exact energies. They are exact for both \( t = 0 \) and \( \lambda = 0 \), as expected, and the crossover from the large to small polaron is captured to a high degree of accuracy. This accuracy is very encouraging, especially since these approximations are so trivial to evaluate. It is also worth pointing out that our work on the Holstein Hamiltonian shows that this accuracy improves in higher-dimensional models, and we believe this to be true here as well. Unfortunately, lack of detailed numerical results in higher dimensions prevents us from confirming this to be the case for models with momentum-dependent coupling.

Two more observations are apparent regarding these results: (i) The energies predicted by MA are lower in energy than the ED results, and (ii) the MA results approach the Lang-Firsov asymptotic limit very slowly.
4.2. Calculating the Green’s Function

Figure 4.1: (a) Ground state energy, (b) Percent difference from ED ground state energy results, (c) Ground state qp weight, and (d) Percent difference from ED qp weight results, as a function of the effective coupling $\lambda$, for $t = 1$, $\Omega / t = 0.5$. The perturbation theory results for both the Holstein and breathing-mode models are shown, and the Holstein result is plotted as a function of $\lambda_H = \lambda / 2$. The ED results are from Ref. [22].

The fact that the energies predicted are below the exact result indicates that the MA approximation is non-variational in the case of the breathing-mode model. This is somewhat surprising, since for the Holstein model we have shown that $\text{MA}^{(0)}$ is variational and that $\text{MA}^{(1)}$ is quasi-variational (in the latter case the Hamiltonian is modified slightly and the approximation is no longer truly variational, see Section 3.2). In any case, the energies found for the Holstein model using the semi-variational $\text{MA}^{(1)}$ and $\text{MA}^{(2)}$ approximations were always higher than the exact numerical results.

The slow asymptotic convergence at large $\lambda$ is due to a different pre-factor of the $\mathcal{O}(t^2)$ perturbational correction. Instead of the correct breathing-mode result from Eq. (1.57):

$$E_B(k) = -\frac{2g^2}{\Omega} + \varepsilon_ke^{-3g^2/\Omega^2} - \frac{\Omega t^2}{g^2} \left[ \frac{1}{3} + \frac{e^{-2g^2/\Omega^2}}{2} \cos(2ka) \right],$$

\[4.24\]
4.2. Calculating the Green’s Function

we have produced the Holstein result from Eq. (1.33), with \( g^2 \) replaced with \( \bar{g}^2 = 2g^2 \) [see the two dashed lines labeled PT in Fig. 4.1(a)]:

\[
E_H(k) = -\frac{2g^2}{\Omega} + \varepsilon_k e^{-2g^2/\Omega^2} - \frac{\Omega t^2}{g^2} \left[ \frac{1}{2} + e^{-2g^2/\Omega^2} \cos(2ka) \right].
\] (4.25)

To understand the origin for both of these facts, we observe that at this level of approximations, almost all dependence on the el-ph coupling \( g_q \) is through its momentum-averaged value \( \bar{g}^2 = (1/N) \sum_q |g_q|^2 \). For the 1D breathing-mode model, this average happens to be the same whether \( g_q \propto i \sin qa^2 \), as is the case here, or whether \( g_q \propto \cos qa^2 \) which would correspond to a coupling of the electron to the sum \( x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}} \) of O-site displacements. In other words, for this model, these simplest versions of the MA approximation register that a phonon cloud is formed at a certain O site, but not whether this leads to a leftwards or rightwards displacement of that site. In the strong coupling limit, one expects clouds to be formed only on the two O sites neighbouring the Cu site that hosts the charge, and to point towards the charge. The effective hopping is related to the overlap of these clouds when the charge hops to a neighbouring site. If the electron hops from \( i \) to \( i+1 \), the phonon cloud at \( i+\frac{1}{2} \) changes from \( \exp \left[ -\frac{g_0 b_{i+\frac{1}{2}}^\dagger}{} \right] |0\rangle \) to \( \exp \left[ \frac{g_0 b_{i+\frac{1}{2}}^\dagger}{} \right] |0\rangle \) and the overlap is small. In the MA approximations, it is precisely this information about left vs. right displacement that is lost (equivalently, one can think of losing the information about relative phases between various contributions to the phonon cloud) and the overlap is unity – both cases have a cloud on the \( i+\frac{1}{2} \) site. This explains the higher polaron mobility, hence the lower energy and slower convergence towards the exact asymptotic value for the MA approximations.

For the Holstein model, this problem does not appear because the el-ph coupling is local. We argue that this problem should also become less serious if the el-ph coupling is longer range, because in that case the relative sign of various displacements is not completely lost in the \( \bar{g}^2 = (1/N) \sum_q |g_q|^2 \) average. Unfortunately, lack of numerical data for lattice models with longer-range el-ph coupling makes it difficult to support this statement. The only comparison we can offer is with the exact solution of a somewhat pathological, infinite-range highly inhomogeneous coupling (HIC) model described in Section 1.6. As shown in Fig. 4.2, for this model both \( \text{MA}^{(0)} \) and \( \text{MA}^{(1)} \) approximations give the correct asymptotic behavior – and this necessitates a correct description of the infinite-range phonon clouds that appear in this model.

To summarize, the MA approximations given by Eqs. (4.15) and (4.22)
Figure 4.2: MA\(^{(0)}\) and MA\(^{(1)}\) predictions of the ground-state energy of the HIC model [73] as compared to its exact solution (black dots). For this model with infinite-range el-ph coupling, the approximations capture the correct asymptotic behavior.

are very easy-to-use, and rather accurate for models with el-ph coupling dependent on the phonon momentum. Given its low dimension and short-range (but not local) el-ph interaction, one would expect the 1D breathing-mode model to be amongst the worst examples for the accuracy of this approximation. However, even here we obtain very decent agreement. As argued, we expect this to improve for higher dimensions and longer-range interactions. The only regime where accuracy is certain to worsen is when \(\Omega/t\) becomes very small (this is a general problem of MA-like approximations, discussed at length in Section 2.2). Therefore, we believe that these approximations are useful for quick guides to relevant energies and other quantities of interest for problems of this type, obtained with minimal analytical and computational effort, yet still reasonably accurate.
4.3 The Variational MA Approximation

In this section, we attempt to remedy the problems pointed out in the previous section and by so doing, to obtain an improved MA approximation for the 1D breathing-mode Hamiltonian. The solution, which is presented in this section, can then be used as a template for any other $g_q$ model. The main idea is to try to formulate an MA approximation which is variational in nature, as the original MA is for the Holstein Hamiltonian.

4.3.1 The MA^{(v,0)} Approximation

Given the good agreement found in the previous section, it is reasonable to use the MA^{(0)} solution as a guidance for what states are most relevant to include. Remember that this solution involved only the fully momentum-average quantities $F_n = (1/N^n) \sum_{q_1,\ldots,q_n} f_n(k,q_1,\ldots,q_n,\omega)$. Using the definition of Eqs. (4.4) and (4.7) and performing the sums over $q_1,\ldots,q_n$ for our specific model with $g_q$ given by Eq. (4.2), we find immediately that

$$F_n \propto \sum_i e^{ikR_i} \langle 0 | c_k \hat{G}(\omega) c_i^{\dagger} (b_i^{\dagger} - b_{i+\frac{1}{2}}^{\dagger} - b_{i+\frac{3}{2}}^{\dagger})^n | 0 \rangle.$$  \hspace{1cm} (4.26)

In other words, only states where there are phonons only on two neighbouring sites contribute to it.

We use this as the criterion for our variational MA approximation. More specifically, when we use the Dyson equation to generate equations of motion, we only keep the terms which have phonons only on two neighbouring sites, and discard any other contribution. Of course, these states have to have a total momentum $k$, or else the matrix element is zero. We find that 3 sets of generalized Green’s functions, sketched in Fig. 4.3, appear when we use this criterion, namely:

$$f_{n,m} = \frac{1}{\sqrt{NG(k,\omega)}} \sum_i e^{ikR_i} \langle 0 | c_k \hat{G}(\omega) c_i^{\dagger} b_i^{\dagger m} b_{i+\frac{1}{2}}^{\dagger m} | 0 \rangle$$  \hspace{1cm} (4.27)

and

$$f_{n,m}^{\pm} = \frac{1}{\sqrt{NG(k,\omega)}} \sum_i e^{ikR_i} \langle 0 | c_k \hat{G}(\omega) c_i^{\dagger \pm 1} b_i^{\dagger m} b_{i+\frac{1}{2}}^{\dagger m} | 0 \rangle.$$  \hspace{1cm} (4.28)

Note: as before, we again do not write explicitly the $k$ and $\omega$ dependence.

These three generalized Green’s functions will provide the variational basis for the improved MA approximation. If one wants to enlarge the variational space by allowing phonons at three or more sites, then one needs
4.3. The Variational MA Approximation

Figure 4.3: The generalized Green’s functions that appear in MA\(^{(0)}\) are sketched in the top picture. They have phonon clouds on two neighbouring sites, with the electron on the central site. For the variational MA approximation, denoted MA\(^{(v,0)}\), we find again such Green’s functions, but also those with the electron to the immediate left or right of the phonon clouds. We name them \(f_{nm}\) and \(f_{nm}^{\pm}\), respectively.

to generate new recurrence relations compatible with that assumption. The total number of generalized Green’s functions increases very fast, though, so we limit ourselves to the assumption of phonon clouds on two neighbouring sites only. As already explained, the MA\(^{(0)}\) accuracy gives us good confidence that this might suffice.

It is straightforward to show that in this notation we can write Eq. (4.8) as

\[
G(k, \omega) = G_0(k, \omega)[1 + g(f_{1,0} - f_{1,1})G(k, \omega)],
\]

and therefore

\[
\Sigma_{\text{MA}^{(v,0)}}(k, \omega) = g(f_{1,0} - f_{1,1}).
\]

The equations of motion consistent with the variational restriction are straight-
where we use the shorthand notation $e$ space Green’s functions, defined as

$$
\begin{align*}
\mathcal{F}_{n,m} &= e_0 \left[ nf_{n-1,m} - nf_{n,m+1} \right] - e_2 e^{-ika} f_{n+1,m} \\
&\quad - e_1 \left[ ne^{-ika} f_{n-1,n-1} - e^{-ika} f_{n+1,n} + e^{-ika} f_{n+1,n+1} - f_{n+1,1} \right].
\end{align*}
$$

(4.31)

For any $1 \leq m \leq n - 1$,

$$
\begin{align*}
\mathcal{F}_{n,m} &= e_0 \left[ (n-m) f_{n-1,m} - mf_{n-1,m-1} + f_{n+1,m} - f_{n+1,m+1} \right] \\
&\quad - e_1 \left[ (n-m) f_{n-1,m} - mf_{n-1,m-1} + f_{n+1,m} - f_{n+1,m+1} \right],
\end{align*}
$$

(4.32)

and

$$
\begin{align*}
\mathcal{F}_{n,n} &= -e_0 \left[ nf_{n-1,n-1} + f_{n+1,n+1} - f_{n+1,1} \right] + e_2 e^{-ika} f_{n+1,1} \\
&\quad + e_1 \left[ ne^{-ika} f_{n-1,0} - e^{-ika} f_{n+1,1} + e^{-ika} f_{n+1,0} - f_{n+1,n} \right].
\end{align*}
$$

(4.33)

Similarly, for any $1 \leq m \leq n - 1$,

$$
\begin{align*}
\mathcal{F}_{n,m}^\pm &= e_1 \left[ (n-m) f_{n-1,m} - mf_{n-1,m-1} + f_{n+1,m} - f_{n+1,m+1} \right] \\
&\quad - \left\{ \begin{array}{c} e_0 \\ e_2 \end{array} \right\} \left[ (n-m) f_{n-1,m} + f_{n+1,m} \right] \\
&\quad + \left\{ \begin{array}{c} e_2 \\ e_0 \end{array} \right\} \left[ mf_{n-1,m} + f_{n+1,m} \right]
\end{align*}
$$

(4.34)

and

$$
\begin{align*}
\mathcal{F}_{n,0}^\pm &= e_1 \left[ nf_{n-1,0} + f_{n+1,0} - f_{n+1,1} \right] + e_0 f_{n+1,1} - e_3 e^{-ika} f_{n+1,n} \\
&\quad - e_2 \left[ ne^{-ika} f_{n-1,n-1} - e^{-ika} f_{n+1,n} + e^{-ika} f_{n+1,n} \right].
\end{align*}
$$

(4.35)

$$
\begin{align*}
\mathcal{F}_{n,n}^\pm &= -e_1 \left[ nf_{n-1,n-1} + f_{n+1,n+1} - f_{n+1,1} \right] - e_0 f_{n+1,n} + e_3 e^{-ika} f_{n+1,n} \\
&\quad + e_2 \left[ ne^{-ika} f_{n-1,0} + e^{-ika} f_{n+1,0} - e^{-ika} f_{n+1,1} \right].
\end{align*}
$$

(4.36)

where we use the shorthand notation $e_j = \tilde{g}_j(\omega - n\Omega)$, and $\tilde{g}_j(\omega)$ are real-space Green’s functions, defined as

$$
\tilde{g}_j(\omega) = \frac{1}{N} \sum_q e^{\pm iq(jn)} G_0(q,\omega) = G_0(j,0,\omega).
$$

(4.37)
4.3. The Variational MA Approximation

The “±” sign in the exponent of Eq. (4.37) is irrelevant because $G_0(q, \omega)$ is even with respect to $q$. The explicit expressions for these and other momentum averaged functions of the free electron propagator are given in Appendix D for a tight-binding dispersion. Note that $f_{n,n}^- = e^{ika} f_{n,0}^-$ and $f_{n,0}^+ = e^{-ika} f_{n,0}$, which is why we do not need to keep them as independent variables.

Equations (4.31)-(4.36) can now be cast in the form

$$v_n = A_n v_{n-1} + B_n v_{n+1},$$

where

$$v_n = \left( f_{n,0}, \ldots, f_{n,n}, f_{n,0}^-, \ldots, f_{n,n-1}^-, f_{n,1}^+, \ldots, f_{n,n}^+ \right)^T \tag{4.38}$$

collects all generalized Green’s functions with a total of $n$ phonons, and the matrices $A_n, B_n$ are straightforward to obtain from the equations above. The solution of this set of recursive equations can then be written as an infinite continued fraction involving products of matrices:

$$v_n = Q_n v_{n-1} = \frac{1}{1 - B_n Q_{n+1} A_n} v_{n-1}. \tag{4.39}$$

Although this is a continued fraction of matrices of increasing size, it is still numerically trivial to evaluate. The dimensions of $A_n$ and $B_n$ are $(3n+1) \times (3n-2)$ and $(3n+1) \times (3n+4)$, respectively, and we determine the truncation level of the continued fraction using the same criteria as discussed in the previous section for the MA$^{(0)}$ approximation. In addition, the $A_n$ and $B_n$ matrices are very sparse, which makes multiplication by them very efficient. Finally, since $v_0 \equiv 1$ by definition, it is straightforward to calculate $\Sigma_{MA^{(v,0)}}(k, \omega)$ from Eq. (4.30) once $v_1 = (f_{1,0}, f_{1,1}, f_{1,0}^-, f_{1,1}^+)^T = Q_1$ is evaluated from Eq. (4.39).

4.3.2 The MA$^{(v,1)}$ Approximation

We can also systematically improve the variational MA approximation to reproduce the electron+one-phonon continuum. It is obvious that this will not be predicted by MA$^{(v,0)}$, since no phonons are allowed to appear far from the main polaronic cloud.

We follow the same approach as before: at the MA$^{(v,1)}$ level we keep all equations involving free electron propagators with $\omega - \Omega$ exactly. In order to work in the enlarged variational space described by Eqs. (4.27) and

\[\text{Such continued fractions involving matrices also appear in the application of MA to the Holstein model with multiple phonon modes, for details see Ref. [15].}\]
4.3. The Variational MA Approximation

(4.28) we need to define the following generalized Green’s functions [these are analogous to the partial momentum averages \( \bar{f}_n(q_1) \) in Eq. (4.21)]:

\[
\bar{f}_{n,m}(q_1) = \frac{g_{q_1}}{\sqrt{NG(k, \omega)}} \sum_i e^{i(k-q_1)R_i} \langle 0 | c_k \hat{G}(\omega) c_i^\dagger b_{i-\frac{1}{2}} m_i b_{i+\frac{1}{2}} | 0 \rangle (4.40)
\]

and

\[
\bar{f}_{n,m}^\pm(q_1) = \frac{g_{q_1}}{\sqrt{NG(k, \omega)}} \sum_i e^{i(k-q_1)R_i} \times \langle 0 | c_k \hat{G}(\omega) c_i^{\pm1} b_{i-\frac{1}{2}} m_i b_{i+\frac{1}{2}} | 0 \rangle . (4.41)
\]

These Green’s functions explicitly contain states with one phonon delocalized away from the main polaronic cloud, i.e. precisely the type of states required to reproduce the polaron+phonon continuum. In this notation Eq. (4.18) can be written as

\[
f_1(q_1) = G_0(k - q_1, \omega - \Omega) \left[ |g_{q_1}|^2 + g(\tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1)) \right], (4.42)
\]

which is an exact equation involving no approximations, as in the MA\(^{(1)}\) case. The equations for \( f_n(q_1) \) with \( n \geq 2 \) can be approximated and solved as in the MA\(^{(v,0)}\) case.

The details of the MA\(^{(v,1)}\) solution are given in Appendix E. The approach is similar to the MA\(^{(1)}\) calculation of Section 4.2.3, and the basic idea is as follows. First, we solve up to the \( n = 2 \) level using the MA\(^{(v,0)}\) result of Eq. (4.39). The exact \( n = 1 \) equation from above can then be cast into a set of recurrence relations similar to Eqs. (4.31)-(4.36), involving the generalized Green’s functions of Eqs. (4.40) and (4.41). After some effort, the resulting system of equations can be reduced to only four equations in four unknowns: \( f_{1,0}^{(1)}, f_{1,1}^{(1)}, f_{1,0}^{-(1)}, f_{1,1}^{+(1)} \). This can be easily solved, and the resulting self-energy is given by

\[
\Sigma_{MA^{(v,1)}}(k, \omega) = g(f_{1,0}^{(1)} - f_{1,1}^{(1)}). (4.43)
\]

Again the MA\(^{(v,1)}\) expression is slightly more involved than the zeroth order approximation because it involves the evaluation of two continued fractions (see Appendix E for details), but it is still numerically trivial to evaluate.

111
4.4 Variational MA Results

We present results for the 1D breathing-mode Hamiltonian using the variational MA approximation. Comparisons are made to ED data [22] where possible.

4.4.1 Ground State Properties

In Fig. 4.4 we plot both the ground state energy and qp weight as a function of the el-ph coupling strength, using the variational MA approximation. The variational MA results show a clear improvement over the MA results shown in Fig. 4.1, and the agreement with the numerical data is excellent. The approximate and exact numerical data results are indistinguishable when plotted over the full parameter range of Fig. 4.4(a). To gain a closer look at the success of our approximation we plot the relative error between the MA results and the ED results in Fig. 4.4(b). Indeed the approximation

Figure 4.4: (a) Ground state energy, (b) Percent difference from ED ground state energy results, (c) Ground state qp weight, and (d) Percent difference from ED qp weight results, as a function of the effective coupling \( \lambda \), for \( t = 1, \Omega/t = 0.5 \). The ED results are from Ref. [22].

4.4 Variational MA Results

We present results for the 1D breathing-mode Hamiltonian using the variational MA approximation. Comparisons are made to ED data [22] where possible.

4.4.1 Ground State Properties

In Fig. 4.4 we plot both the ground state energy and qp weight as a function of the el-ph coupling strength, using the variational MA approximation. The variational MA results show a clear improvement over the MA results shown in Fig. 4.1, and the agreement with the numerical data is excellent. The approximate and exact numerical data results are indistinguishable when plotted over the full parameter range of Fig. 4.4(a). To gain a closer look at the success of our approximation we plot the relative error between the MA results and the ED results in Fig. 4.4(b). Indeed the approximation
4.4. Variational MA Results

is giving extremely close agreement with the numerical results, with less than 0.3% relative error for both $\text{MA}^{(v,0)}$ and $\text{MA}^{(v,1)}$. The largest errors occur at intermediate couplings, as expected because the MA approximation is exact in both the zero coupling and zero bandwidth limits. We also show a comparison of the quasiparticle weight calculated using variational MA with the ED result in Fig. 4.4(c). Again, when plotted over the full parameter range, the results are nearly indistinguishable. A look at the percent difference indicates that the relative error is less than 2.5%. This is quite a remarkable fact considering that the qp weight contains information on the nature of the eigenstates, something that is rarely obtained accurately when using approximate methods.

It is clear from Fig. 4.4 that the variational MA approximation cures all of the shortcomings of the simple MA generalization discussed in Section 4.2.4. A comparison of the variational MA and ED results shows that the MA energies are slightly higher than the exact numerical results, as expected from a variational method, and that the asymptotic behavior predicted from both the perturbational theory and ED result is reproduced. These successes were expected because the variational MA approximation was designed precisely to remedy these problems, but it is still very encouraging to see that the physical picture described and used to motivate the variational MA approximation in Section 4.3 was indeed correct.

In Fig. 4.5(a) we plot the average number of phonons in the cloud, $N_{\text{ph}}$, as calculated from the Hellmann-Feynman theorem (see Section 2.3). Since there is no numerical data available for this quantity we compare our findings to the standard Rayleigh-Schrödinger (RS) perturbation theory result at small couplings [see Eq. (1.50)], and to the strong-coupling perturbation theory result for larger couplings [see Eq. (1.57)]. The approximation reproduces both asymptotic limits to a high degree of accuracy, as should be expected based on the success of the ground state energies and qp weights shown in Fig. 4.4. In Fig. 4.5 we also plot the effective mass as a function of the coupling strength. Again the agreement with the numerical data is excellent, as confirmed by the relative errors shown in Fig. 4.5(c).

4.4.2 The Polaron Band

With our analytical expression for the self-energy we can also calculate momentum-dependent results. In Fig. 4.6(a) we plot the lowest energy state for momenta $0 \leq k \leq \pi$ and compare our results to the available numerical data. We again find that the variational MA approximation is highly accurate. Because the polaron dispersion is relatively flat for the in-
4.4. Variational MA Results

Figure 4.5: (a) Average number of phonons and (b) effective mass, (c) Percent difference from ED effective mass results, as a function of the effective coupling $\lambda$, for $t = 1, \Omega/t = 0.5$. The ED results are from Ref. [22].
4.4. Variational MA Results

Figure 4.6: (a) Polaron dispersion $E_k$, (b) Percent difference from ED ground state energy results, (c) quasiparticle weight $Z_k$, and (d) Percent difference from ED qp weight results. Results are shown for $t = 1, \Omega/t = 0.5$, and $\lambda = 1.071502$. The ED results are from Ref. [22].

intermediate coupling strength of $\lambda = 1.071502$ shown here, the energy range of the plot is quite narrow and we can clearly discern the difference between the approximate and numerical results. However, by looking at the relative error in Fig. 4.6(b), we demonstrate that the accuracy of the variational MA result is again very good, coming well within 0.5% relative error of the numerical result, showing that the variational MA approximation is accurate for all momenta $k$.

The reason for the non-monotonic polaron dispersion was discussed in Section 1.5, and is due to a larger effective 2nd nearest neighbour hopping than the effective nearest neighbour hopping of the polaron [see Eq. (1.57)]. This is a direct consequence of the structure of the polaronic cloud. In fact, at the MA$(0)$ and MA$(1)$ levels of approximation we find a monotonically increasing dispersion, much like the Holstein case, even though the self-energy is explicitly $k$-dependent for all MA$(n)$ results. This is not surprising in light of the shortcomings of MA$(n)$ discussed in Section 4.2.4, where we showed that MA$(n)$ does not correctly reproduce the second-order strong coupling
4.4. Variational MA Results

...perturbation theory result, precisely the origin of the non-monotonic dispersion seen here. In any case, the MA\(^{(v,0)}\) and MA\(^{(v,1)}\) self-energies are strongly \(k\)-dependent, in agreement with the numerical data shown in Fig. 4.6(a).

In Figs. 4.6(c) and 4.6(d) we also plot the qp weight and its relative error as a function of momenta. We again find good agreement. As explained, the agreement is expected to improve for both smaller and larger \(\lambda\) values, where our approximations become asymptotically exact.

4.4.3 Higher Energy Properties

Lastly, we consider the high energy properties of the 1D breathing-mode Hamiltonian using MA\(^{(v,0)}\) and MA\(^{(v,1)}\). In Fig. 4.7 we compare our predicted spectral weights to available numerical data. When plotted on a linear axis, the results are essentially indistinguishable, especially for MA\(^{(v,1)}\). To gain a better view we display the same plots on a logarithmic scale in Fig. 4.7(b). There are a few interesting features to note. At the MA\(^{(0)}\) level (red dashed line) the majority of the spectral weight is found to be in the correct location, however, the electron+phonon continuum is completely absent for the parameters shown, as expected. At the MA\(^{(v,1)}\) level the continuum is reproduced in the expected location, in very good agreement with the ED prediction. Furthermore, the finite size effects responsible for the sharp peaks in the continuum of the ED result are absent from the MA\(^{(v,1)}\) data. As a guide illustrating the correct location for the polaron+phonon continuum we have added the vertical blue dashed lines to denote \(E_0\) and \(E_0 + \Omega\) for the MA\(^{(v,1)}\) ground state energy. The lower edge is located correctly, however especially for smaller \(\lambda\), MA predicts a somewhat wider continuum than ED. Given the very limited availability of numerical results of this type, we do not know if this discrepancy is due to truncation approximations in the ED solution, or is due to inaccuracies of the MA approximations.

In Fig. 4.8 we plot the spectral weight for a fixed coupling strength and vary the momentum \(k\). Again on the linear scale the results are very encouraging. The MA result, particularly MA\(^{(v,1)}\) predicts the correct location for the spectral weight over the entire energy range. However, a logarithmic scale plot does reveal a notable shortcoming of the approximation. For \(k = 0\) the agreement with the numerical data is excellent and the continuum is located at the expected \(E_0 + \Omega\). For \(k = 0.25\pi\) and \(k = 0.5\pi\) the agreement is still good, but as we begin to approach the band edge we see a significant deviation of the MA result from the ED result, with the MA continuum coming to much too low energies. This feature has extremely little
4.4. Variational MA Results

Figure 4.7: (a) $A(k = 0, \omega)$ vs. $\omega$ and (b) $\ln A(k = 0, \omega)$ for $t = 1, \Omega = 0.5, \eta = 0.004$, and $\lambda = 0.75, 1.0, 1.25, 1.5$; The vertical blue dashed lines denote $E_0$ and $E_0 + \Omega$ using $\text{MA}^{(v,1)}$. The ED results are from Ref. [22].
Figure 4.8: (a) $A(k, \omega)$ and (b) $\ln A(k, \omega)$ vs. $\omega$ for $k/\pi = 0, 0.25, 0.5, 0.75, 1$, $t = 1, \Omega = 0.5, \lambda = 1.25, \eta = 0.004$; The ED results are from Ref. [22].
spectral weight [see Fig. 4.8(a)] and is extremely unlikely to be detectable in any experimental realization, however it is in stark disagreement with the expected result confirmed by the ED data. We do not currently understand this failure at higher $k$ values.

Finally, in Fig. 4.9 we show more detailed plots of the evolution of the spectral weight for $k = 0$ from zero coupling to large coupling, for a full range of energies. The two panels compare the MA$^{(v,0)}$ (a) and MA$^{(v,1)}$ (b) solutions. The one obvious difference is the location of the polaron+phonon continuum, which is incorrectly located around $-2t$ for MA$^{(v,0)}$, whereas it always starts at $\Omega$ above $E_0$ for MA$^{(v,1)}$. Aside from this feature with rather little weight, all the features which contribute significantly to the spectral weight are in good agreement.

Plots of this nature would require extensive numerical computational time if one tried to generate them using “exact” computational methods. For example, achieving convergence for the ED results at larger $\lambda$ required inclusion of billions of states, in some cases, in the truncated Hilbert space. Diagonalizing such large (even though sparse) matrices is not a trivial task. By comparison, because we have an analytical expression for the self-energy, these detailed MA plots covering many energies and couplings take just seconds, at most minutes, to generate. In fact, MA is much faster even than SCBA, which for this model requires many numerical integrals over the Brillouin zone (and is of questionable use for medium and large couplings).

4.5 Conclusion

In summary, we have presented a way to generalize the momentum average approximation to el-ph coupling models with momentum-dependent couplings. The approximation still sums all of the self-energy diagrams, albeit with each diagram approximated in such a way that the full sum can be performed, and it is exact in both the zero coupling and zero bandwidth limits. As in the application of MA to the Holstein model, the approximation is analytical and easy-to-use, and gives highly accurate results over the entire parameter space. In this chapter we have actually presented two different generalizations of MA, a straightforward generalization that can easily be applied to any model with a momentum-dependent coupling with minimal effort, and a more specific (and extremely accurate) generalization that takes the details of a given model into account and enlarges the variational subspace accordingly. We have used the 1D breathing-mode Hamiltonian as an example to gauge the accuracy of both of these MA generalizations.
4.5. Conclusion

Figure 4.9: Spectral weight $A(0, \omega)$ vs. $\omega$ in 1D using (a) $\text{MA}^{(v,0)}$ and (b) $\text{MA}^{(v,1)}$. The results are shown for $t = 1, \Omega = 0.5, \eta = 0.01$, and $\lambda$ varying from 0 to 2. Curves corresponding to $\lambda = 0.3, 0.6, 1.2$ and $1.8$ are highlighted in red.
4.5. Conclusion

While the straightforward extension of MA gave fairly accurate results, it was non-variational and it approached the zero-bandwidth asymptotic limit very slowly. The variational MA approximation remedied both of these problems and produced extremely accurate results, coming well within 0.3% error of the available numerical results for the ground state energies of the breathing-mode polaron. We also showed that MA can be systematically improved in both cases, leading to higher accuracy and the correct location for the electron+phonon continuum.

The successful generalization of MA to this much broader class of models is very encouraging. Numerical studies of models as complicated as even the 1D breathing-mode Hamiltonian are very intensive, and the MA approximation provides a quick and easy-to-use way to gain an understanding of these more realistic models without having to do a detailed numerical analysis from the start. We hope that this tool will be useful for probing more realistic models of interesting physical systems.
Chapter 5

The Effect of a Surface on a Holstein Polaron

All the work so far has been focused on calculations of bulk properties such as densities of states and spectral weights. However, in practice many materials, and in particular the layered compounds, are being investigated extensively using surface-sensitive spectroscopies, such as surface tunneling microscopy (STM) and Angle-Resolved Photoemission Spectroscopy (ARPES). Although the former probes only the surface layer while the latter probes the first few layers, they are often interpreted in terms of bulk densities of states or spectral weights, especially for surfaces which are not expected to bind surface states. An example of a surface having el-ph coupling properties significantly different the bulk is that of Be(0001). Two independent groups have shown that the el-ph coupling constant for the \( \bar{\Gamma} \) surface state is several times larger than the bulk value near the Fermi energy \([2, 74–76]\). Electron-phonon coupling plays an important role in other surfaces as well, for e.g. Mo(110) \([86]\). A review can be found in Ref. \([23]\).

In this chapter we study the problem of a Holstein polaron near a surface, and show that its properties are considerably changed from the expected bulk results. For example, because of the particle-boson coupling, surface states can appear even for surfaces where they are not expected. These results clearly demonstrate that interpretation of the data from surface spectroscopies in terms of bulk properties is not necessarily simple or warranted.

To study the effect of a surface on Holstein polarons we further generalize the MA approximation to account for broken translational symmetry. This inhomogeneous MA (IMA) approximation continues to be accurate over the entire range of el-ph coupling strengths, as was the case with previous versions of MA, as it takes into account the broken symmetry without any further approximations. We also note that the IMA approximation has been co-developed for understanding the effects of disorder on polaron properties \([87]\). In this study involving disorder IMA results for a 1D Holstein model with an impurity at the origin are compared to highly accurate but rather
expensive numerical simulations based on the Diagrammatic Monte Carlo method [51, 87]. There it is shown that the high accuracy of MA is carried over to IMA. This is as expected, as we will show below, because IMA is based on the same physical arguments as the original MA approximations and it obeys the same number of spectral weight sum rules.

This chapter is organized as follows. In Sections 5.1 and 5.2 we describe our model and derive the IMA approximation. In Section 5.3 we present our results, showing the effects of a surface on the spectra of Holstein polarons, and Section 5.4 contains our summary and conclusions.

5.1 The Model of the Surface

We first describe how we model the surface and calculate the associated Green’s functions in the absence of el-ph coupling. For simplicity, we consider crystal-vacuum interfaces and do not allow for the polaron wave function to spill into the vacuum. This approximation, which implies an infinite work function, can be relaxed and one could in principle also study interfaces where two different kinds of materials (with different band structure and/or different particle-boson coupling) are in contact, but we will not discuss these more complicated cases here. To demonstrate the versatility of this method, we also allow for addition of a surface potential, which can be made negative or positive in order to control the tendency to bind surface states. We assume this potential to be uniform, so as to not break translational invariance in directions parallel to the surface. However, localized impurity potentials can also be added and treated within this approach, allowing for realistic modeling of STM spectra in the vicinity of various adsorbed atoms. Since we treat the effect of the surface without any further approximations, we expect the accuracy of these results to be as high as for all other MA-based results. To our knowledge these are the first available results for such problems.

To begin, we show how to calculate the Green’s function of the system with the surface, in the absence of el-ph coupling. We first solve explicitly the problem for a semi-infinite 1D chain and then generalize to higher dimensions. For an \( N \)-site chain with periodic boundary conditions described by nearest-neighbour hopping we have

\[
\hat{H}_0 = -t \sum_{\langle i,j \rangle} \left( c_i^\dagger c_j + h.c. \right).
\]  (5.1)
5.1. The Model of the Surface

Its Green’s function is well-known [80]:

\[ G_0(i,j,\omega) = \langle 0 | c_i \hat{G}_0(\omega) c_j^\dagger | 0 \rangle = \frac{1}{N} \sum_k \frac{e^{ik(R_i-R_j)}}{\omega - \varepsilon_k + i\eta}, \]  \hspace{1cm} (5.2)

where the sum is over the Brillouin zone (BZ) \( k \in (-\frac{\pi}{a}, \frac{\pi}{a}] \), and the sum becomes an integral in the limit \( N \to \infty \). For nearest-neighbour hopping the electron dispersion is given by the usual \( \varepsilon_k = -2t \cos(ka) \) in 1D, where \( a \) is lattice spacing, but our results can be applied for any dispersion.

To calculate the free electron Green’s function on a semi-infinite chain that begins at site 0, \( G_c(i,j,\omega), \forall i, j \geq 0 \), we “cut” the infinite chain between sites \( i = -1 \) and \( i = 0 \) by adding a term

\[ \hat{V}_c = +t \left( c_{-1}^\dagger c_0 + c_0^\dagger c_{-1} \right) \]  \hspace{1cm} (5.3)

to the free electron Hamiltonian \( \hat{H}_0 \), such that \( \hat{H}_c = \hat{H}_0 + \hat{V}_c \) does not allow hopping to the left of site 0. We then apply Dyson’s identity

\[ \hat{G}_c(\omega) = \hat{G}_0(\omega) + \hat{G}_0(\omega) \hat{V}_c \hat{G}_c(\omega) \]  \hspace{1cm} (5.4)

which gives

\[ G_c(i,j,\omega) = G_0(i,j,\omega) + tG_0(i,-1,\omega)G_c(0,j,\omega). \]  \hspace{1cm} (5.5)

This can be easily solved to give

\[ G_c(i,j,\omega) = G_0(i,j,\omega) + t \frac{G_0(i,-1,\omega)G_0(0,j,\omega)}{1 - tG_0(0,-1,\omega)}. \]  \hspace{1cm} (5.6)

Similar steps are used to account for a surface potential at site 0, which we model by an additional potential

\[ \hat{U}_0 = -U c_0^\dagger c_0. \]  \hspace{1cm} (5.7)

Writing the total Hamiltonian of the chain as \( \hat{H}_s = \hat{H}_c + \hat{U}_0 \), we can again apply Dyson’s identity, \( \hat{G}_s(\omega) = \hat{G}_c(\omega) + \hat{G}_c(\omega) \hat{U}_0 \hat{G}_s(\omega) \), giving

\[ G_s(i,j,\omega) = G_c(i,j,\omega) - U G_c(i,0,\omega) G_s(0,j,\omega). \]  \hspace{1cm} (5.8)

As before this is easily solved, and the result is

\[ G_s(i,j,\omega) = G_c(i,j,\omega) - U \frac{G_c(i,0,\omega) G_c(0,j,\omega)}{1 + U G_c(0,0,\omega)}. \]  \hspace{1cm} (5.9)
For more complicated surface potentials one can calculate $G_s(i,j,\omega)$ in a similar fashion. Because this is a single electron Hamiltonian it can be diagonalized exactly, and therefore $G_s$ is always known exactly.

The above results can now be generalized to higher dimensions. We consider a $d$-dimensional simple cubic crystal with tight-binding nearest-neighbour hopping, cut along a (100)-like surface, so that only sites with $i_x \geq 0$ are in the crystal. We discuss this simplest possible surface for illustrative purposes, but any others can be treated similarly [88]. Because of symmetry to in-plane translations, it is convenient to partially Fourier transform and use operators $c_{i,k} (c_{i,k}^\dagger)$, which annihilate (create) a particle with in-plane momentum $\mathbf{k}$ from (in) the layer $i \geq 0$. In this notation the components of the full Hamiltonian become

$$\hat{H}_0 = \sum_{\mathbf{k}} \left[ \sum_i \varepsilon_{||,\mathbf{k}} c_{i,\mathbf{k}}^\dagger c_{i,\mathbf{k}} - t \sum_{(i,j)} \left( c_{i,\mathbf{k}}^\dagger c_{j,\mathbf{k}} + \text{h.c.} \right) \right], \quad (5.10)$$

$$\hat{V}_c = +t \sum_{\mathbf{k}} \left( c_{-1,\mathbf{k}}^\dagger c_{0,\mathbf{k}} + c_{0,\mathbf{k}}^\dagger c_{-1,\mathbf{k}} \right), \quad (5.11)$$

and

$$\hat{U}_0 = -U \sum_{\mathbf{k}} c_{0,\mathbf{k}}^\dagger c_{0,\mathbf{k}}; \quad (5.12)$$

where $\varepsilon_{||,\mathbf{k}} = -2t_{||} \sum_{i=1}^{d-1} \cos(k_i a)$ is the kinetic energy associated with the in-plane momentum. Since $\mathbf{k}$ is a good quantum number we have

$$\langle 0 | c_{i,\mathbf{k}}^\dagger \hat{G}_s(\omega) c_{j,\mathbf{k}} | 0 \rangle = \delta_{\mathbf{k},\mathbf{k'}} G_s(\mathbf{k},i,j,\omega) \quad (5.13)$$

and the problem factorizes into a sum of 1D problems, one corresponding to each in-plane momentum $\mathbf{k}$. As a result

$$G_s(\mathbf{k},i,j,\omega) = G_s(i,j,\omega - \varepsilon_{||,\mathbf{k}}), \quad (5.14)$$

where $G_s(i,j,\omega)$ is the 1D solution of Eq. (5.9). Note that because of the clear distinction between where the in-plane hopping $t_{||}$ appears in $\varepsilon_{||,\mathbf{k}}$ and where the inter-layer hopping $t$ appears in $G_s$, it is trivial to treat layered compounds with anisotropic hopping. We will study the effects of this anisotropy in Section 5.3.2.

We also note that (100)-like cuts do not bind surface states in the absence of a surface potential. Indeed, in this case $G_s = G_c$ [see Eq. (5.9)], which has no poles outside the continuum $\omega \in [-2t,2t]$, irrespective of the value of $t$ (we have taken $t = t_{||}$). Such poles appear for a surface potential $U \geq t$,}
5.1. The Model of the Surface

Figure 5.1: The 2D spectral weight $A(k, i, i, \omega)$ vs. $\omega$ in the absence of el-ph coupling, for the first 5 surface layers and bulk. The results are obtained from Eq. (5.14) with $\eta = 0.004t$ and $t = t_\parallel$.

indicating the formation of Tamm surface states [89]. This is true for any dimension $d$.

We define the spectral function in the absence of any el-ph coupling as

$$A_s(k, i, i, \omega) = -\frac{1}{\pi} \text{Im} G_s(k, i, i, \omega),$$

where $k$ is the in-plane momentum of the electron and $G_s(k, i, i, \omega)$ is given by Eq. (5.14). This is the spectral weight measured by Angle-Resolved Inverse Photoemission Spectroscopy (ARIPES) if an electron with in-plane momentum $k$ is injected in layer $i$. We also point out that ARPES in a full band would see the same thing, but with “flipped” energies, i.e. $A(k, i, i, -\omega)$.

In Fig. 5.1 we show the 2D spectral weight $A_s(k, i, i, \omega)$ for the first
several layers $i$ with varying surface potential strength $U$. By increasing $U$ we show that a bound state just starts to form at $U/t = 1$, and that there is a clearly a bound state formed for $U/t = 2$. The effect of this surface potential is seen several layers into the bulk but it is found to decrease exponentially with increasing $i$. We find similar behaviour in other dimensions.

5.2 The Holstein Model with a Surface

Including el-ph interactions, the full Hamiltonian for a $d$-dimensional system with a (100)-like cut is given by

$$\hat{H} = \hat{H}_s + \hat{H}_{\text{ph}} + \hat{V}_{\text{el-ph}},$$

(5.16)

where $\hat{H}_s$ is the sum of the terms in Eqs. (5.10)-(5.12) already discussed above, and $\hat{H}_{\text{ph}}$ describes a branch of Einstein optical phonons of energy $\Omega$:

$$\hat{H}_{\text{ph}} = \Omega \sum_{i \geq 0} \sum_{q} b^\dagger_{i,q} b_{i,q},$$

(5.17)

where $b^\dagger_{i,q}$ and $b_{i,q}$ describe the creation and annihilation operators, respectively, for a phonon in plane $i$ with in-plane momentum $q$. The electron-phonon coupling is described by

$$\hat{V}_{\text{el-ph}} = \frac{g}{\sqrt{N||}} \sum_{i \geq 0} \sum_{k,q} c^\dagger_{i,k-q} c_{i,k} \left( b^\dagger_{i,q} + b_{i,-q} \right),$$

(5.18)

which is the usual Holstein coupling, partially Fourier transformed to account for the in-plane momentum of the polaron. $N|| \to \infty$ is the number of unit cells in the surface edge (if $d = 2$), or the surface layer (if $d = 3$), and the $k, q$ sums are over the in-plane BZ. For a chain ($d = 1$) these terms disappear.

In general one expects phonon modes to soften near a surface. Although we do not consider such cases here, this can be addressed within the formalism presented below. We plan to study this effect in a future work.

5.2.1 Exact Solution

Since the total Hamiltonian is invariant to translations parallel to the surface, it is convenient to work with the partially Fourier transformed operators. In this notation, the polaron Green’s functions can be written as

$$G(k, i, j, \omega) = \langle 0 | c_{i,k} \hat{G}(\omega) c^\dagger_{j,k} | 0 \rangle = \sum_{\alpha} \frac{\langle 0 | c_{i,k} | \alpha \rangle \langle \alpha | c^\dagger_{j,k} | 0 \rangle}{\omega - E^\alpha + i\eta},$$

(5.19)
5.2. The Holstein Model with a Surface

where \( \hat{G}(\omega) \equiv [\omega - H + i\eta]^{-1} \), and \( H \) is the full Hamiltonian of Eq. (5.16).

As we did for the Holstein model in Section 2.1.1 and the more general el-ph coupling model in Section 4.2.1, we begin by deriving the exact equations of motion for the Green’s function of the polaron. We proceed through repeated use of Dyson’s identity: 

\[
\hat{G}(\omega) = \hat{G}_0(\omega) + \hat{G}(\omega) \hat{V}_{\text{el-ph}} \hat{G}_s(\omega).
\]

Applying the identity once, we obtain

\[
G(k, i, j, \omega) = G_s(k, i, j, \omega) + \frac{g}{\sqrt{N}} \sum_{j_1, q_1} F_1(k, q_1, i, j_1, \omega) G_s(k, j_1, j, \omega),
\]

(5.20)

where we have defined the generalized Green’s functions

\[
F_n(k, q_1, \ldots, q_n, i, j_1, \ldots, j_n, \omega) = \langle 0| c_{i, \mathbf{k}} \hat{G}(\omega) c_{j_n, \mathbf{k} - \mathbf{q}_n} b_{j_1, q_1}^\dagger \cdots b_{j_n, q_n}^\dagger |0 \rangle.
\]

(5.21)

In principle one can apply Dyson’s identity to Eq. (5.21) and generate the exact higher order equations of motion. However, these expressions are very lengthy and not particularly enlightening. Since we know that we will need to apply the MA approximation to solve the resulting exact equations of motion, we can simplify things considerably from the onset via the following steps. In real space we know that MA only retains states with all phonons on a single site. Therefore, we certainly need not allow for phonons to be spread over several layers, as in Eq. (5.21), because we will ignore such contributions when we apply MA later on. This allows us to write Eq. (5.21) in the following (approximated) form:

\[
F_n(k, q_1, \ldots, q_n, i, j, \omega) = \langle 0| c_{i, \mathbf{k}} \hat{G}(\omega) c_{j_n, \mathbf{k} - \mathbf{q}_n} b_{j_1, q_1}^\dagger \cdots b_{j_n, q_n}^\dagger |0 \rangle.
\]

(5.22)

Applying Dyson’s identity to this simplified expression, we find that for \( n \geq 1 \)

\[
F_n(q_1, \ldots, q_n, i, j) = \frac{g}{\sqrt{N}} \sum_{j \neq j'} G_s(k - q_{j'j}, \omega - n\Omega) \times \langle 0| c_{i, \mathbf{k}} \hat{G}(\omega) \sum_{q_{n+1}} c_{j_{n+1} q_{n+1}}^\dagger \cdots b_{j_n, q_n}^\dagger \cdots b_{j_1, q_1}^\dagger b_{j'j}^\dagger b_{j'j}^\dagger \cdots b_{j_1, q_1}^\dagger |0 \rangle
\]

\[
+ \frac{g}{\sqrt{N}} G_s(k - q_{j'j}, \omega - n\Omega) \sum_{i=1}^n F_{n-1}(\ldots, q_{i-1}, q_{i+1}, \ldots, i, j)
\]

\[
+ \sum_{q_{n+1}} F_{n+1}(q_1, \ldots, q_{n+1}, i, j),
\]

(5.23)
5.2. The Holstein Model with a Surface

where \( q_T = \sum_{i=1}^{n} q_i \), and we have also introduced the shorthand notation \( F_n(k, q_1, \ldots, q_n, i, j, \omega) \equiv F_n(q_1, \ldots, q_n, i, j) \) (i.e., the \( k \) and \( \omega \) dependence of these functions is implicitly assumed from now on). Equation (5.23) shows that \( F_n \) can be written in terms of functions with \( n - 1 \) and \( n + 1 \) phonons only. Also note that \( F_0(k, i, j, \omega) = G(k, i, j, \omega) \).

5.2.2 The IMA Approximation

At the IMA\(^{(0)}\) level of approximation we set \( G_s(k - q_T, j', j, \omega - n\Omega) \to 0 \) if \( j \neq j' \) and \( n > 0 \) in all of these equations, allowing us to write Eq. (5.23) in terms of \( F_n \) functions only. This can be justified as follows. In Section 2.2.2 we demonstrated that the MA approximation can be interpreted as keeping only the diagonal contributions of the free propagator \( G_0(i, j, \omega - n\Omega) \) in real space, for \( n > 0 \), which is precisely what we have done here. Therefore, we expect the accuracy of this approximation to be the same as for the bulk Holstein case. As before, we argue that this is a good approximation, at least for low energies, because the free propagator decreases exponentially with the distance \(|i - j|\), and this exponential decrease is steeper for larger \( n \). It is also easy to check that this approximation becomes exact for both \( t = 0 \) and \( g = 0 \), and it is accurate at higher energies because the MA approximation obeys multiple sum rules (see Section 2.2).

For \( d > 1 \) we must also use the MA approximation to treat the in-plane properties of the polaron. This procedure is analogous to the bulk Holstein case. Observing from Eq. (5.23) that all \( F_n \) are proportional to \( G \), we can recast our equations into a more convenient form by defining

\[
f_n(q_1, \ldots, q_n, i, j) = \frac{N^{n/2}g^n F_n(q_1, \ldots, q_n, i, j)}{G(k, \omega)}. \tag{5.24}
\]

In this notation, and with the approximation made above, Eqs. (5.20) and (5.23) reduce to

\[
G(k, i, j, \omega) = G_0(k, i, j, \omega) \left[ 1 + \frac{1}{N^{d/2}} \sum_{q_1, j'} f_1(q_1, i, j') G(k, j', j, \omega) \right]. \tag{5.25}
\]
5.2. The Holstein Model with a Surface

\[ f_n(q_1, \ldots, q_n, i, j) = G_s(k - q_T, j, j, \omega - n\Omega) \]

\[ \times \left[ g^2 \sum_{i=1}^{n} f_{n-1}(\ldots, q_{n-1}, q_{n+1}, \ldots, i, j) + \frac{1}{N||_{q_{n+1}}} \sum_{q_{n+1}} f_{n+1}(q_1, \ldots, q_n, i, j) \right], \]

(5.26)

respectively. We then define the momentum averaged Green’s functions

\[ F_n(i, j) = \frac{1}{N||} \sum_{q_1, \ldots, q_n} f_n(q_1, \ldots, q_n, i, j), \]

(5.27)

and for \( n > 0 \) we replace \( G_s(k - q_T, j, j, \omega - n\Omega) \) by their momentum averages (this is the usual MA approximation):

\[ \bar{G}_s(j, j, \omega - n\Omega) = \frac{1}{N||} \sum_q G_s(k - q, j, j, \omega - n\Omega). \]

(5.28)

This allows Eq. (5.26) to be written as

\[ F_n(i, j) = \bar{G}_s(j, j, \omega - n\Omega) \left[ n g^2 F_{n-1}(i, j) + F_{n+1}(i, j) \right], \]

(5.29)

which has a solution in terms of continued fractions [see Eq. (2.13)]:

\[ F_n(i, j) = g^2 A_n(j, \omega) F_{n-1}(i, j), \]

(5.30)

where

\[ A_n(j, \omega) = \frac{n G_s(j, j, \omega - n\Omega)}{1 - g^2 \bar{G}_s(j, j, \omega - n\Omega) A_{n+1}(j, \omega)}. \]

(5.31)

This result can then be inserted in Eq. (5.25) to obtain the IMA\(^{(0)}\) expression for the Green’s function:

\[ G_{IMA\,(0)}(k, i, j, \omega) = G_s(k, i, j, \omega) \]

\[ + \sum_{j' \geq 0} G_{IMA\,(0)}(k, j, j', \omega) \Sigma_{IMA\,(0)}(j', \omega) G_s(k, j', j, \omega), \]

(5.32)

where \( \Sigma_{IMA\,(0)}(j, \omega) = g^2 A_1(j, \omega). \)

In principle Eq. (5.32) is the final IMA\(^{(0)}\) expression for the surface Green’s function, but convergence can be obtained more readily by utilizing the bulk MA result for the Holstein polaron. At far distances from the
5.2. The Holstein Model with a Surface

surface the effect of the surface and potential decay to zero. Therefore, the IMA approximation becomes equivalent to the MA approximation as $j \to \infty$:

$$A_n(\omega) = A_n(j, \omega)|_{j \to \infty},$$  \hspace{1cm} (5.33)

where $A_n(\omega)$ is the bulk Holstein result is given by $\Sigma_{\text{MA}}^{(0)} = g^2 A_1(\omega)$ [see Eq. (2.25)]. We can then define an effective interaction

$$v_0(j, \omega) = \Sigma_{\text{IMA}}^{(0)}(j, \omega) - \Sigma_{\text{MA}}^{(0)}(\omega)$$  \hspace{1cm} (5.34)

as the difference between the bulk and surface self-energies to rewrite Eq. (5.32) as

$$G_{\text{IMA}}^{(0)}(k, i, j, \omega) = G_s(k, i, j, \tilde{\omega}) + \sum_{j' \geq 0} G_{\text{IMA}}^{(0)}(k, j', j, \omega) v_0(j', \omega) G_s(k, j, j', \tilde{\omega}),$$  \hspace{1cm} (5.35)

where $\tilde{\omega} = \omega - \Sigma_{\text{MA}}^{(0)}$. This equation is very efficient to evaluate because the bulk properties of the Holstein polaron have been absorbed into the free propagators, and the effective interaction $v_0(j, \omega)$ goes to zero rapidly for increasing $j$. We typically require a very small cutoff for convergence, on the order of $j \leq 5$. However, a cutoff of 0 or 1 is not sufficient for convergence, showing that $v_0(j, \omega)$ is spread over a few sites near the surface.

Equation (5.35) reveals that the el-ph interaction has a very interesting effect on the electron. If the solution were just $G(k, i, j, \omega) = G_s(k, i, j, \tilde{\omega})$ we would conclude that the polaron quasiparticle feels the effect of surface and its on-site potential at the surface only through having its energy renormalized by the self-energy, which accounts for the renormalization of the polaron effective mass, etc. through $G_s$. However, the second term reveals that the surface potential is renormalized via

$$\hat{V}_s \rightarrow \hat{V}_s + \sum_j v_0(j, \omega) c_j^\dagger c_j,$$  \hspace{1cm} (5.36)

in other words the el-ph interaction also results in an additional surface potential with retardation effects through its $\omega$-dependence. This will be discussed more in Section 5.3.

Since the effects of the surface and the on-site potential are taken into account exactly through $G_s(k, i, j, \omega)$, we expect the accuracy and success of IMA to be the same as that for the bulk Holstein model, and we also expect the approximation to have similar improvements at the IMA(1) level, which we derive now.
At the IMA\(^{(1)}\) level, we allow for states where an extra phonon can appear away from the main cloud, thus producing the polaron+one-phonon continuum that must occur at \(E_{GS} + \Omega\). The details of the derivation of the IMA\(^{(1)}\) result are outlined in Appendix F. The basic idea is the same as for the bulk Holstein model, and can be summarized as follows. At this level of approximation we replace \(G_s(k - q_{T\ell}, j', j, \omega - n\Omega) \to 0\) for \(j \neq j'\) and \(n \geq 2\). In other words, we treat the equation for \(F_1\) in Eq. (5.23) exactly, and make approximations for functions with 2 or more phonons. The resulting system of equations can be solved and cast into a form similar to Eq. (5.35) (see Appendix F for details):

\[
G_{IMA\,(1)}(k, i, j, \omega) = G_s(k, i, j, \tilde{\omega}) + \sum_{j' \geq 0} G_{IMA\,(1)}(k, j, j', \omega)v_1(j', \omega)G_s(k, j', j, \tilde{\omega}), \quad (5.37)
\]

where \(\tilde{\omega} = \omega - \Sigma_{IMA\,(1)}\), and \(\Sigma_{IMA\,(1)}\) is the MA\(^{(1)}\) bulk self-energy defined in Eq. (3.7). The effective interaction \(v_1(j, \omega)\) has a more complicated (but more accurate) expression at the IMA\(^{(1)}\) level:

\[
v_1(j, \omega) = \Sigma_{IMA\,(1)}(j, \omega) - \Sigma_{MA\,(1)}(\omega), \quad (5.38)
\]

where

\[
\Sigma_{IMA\,(1)}(j, \omega) = \frac{g^2 \bar{x}_{jj}(\omega)}{1 - g^2 \bar{x}_{jj}(\omega)[A_2(j, \omega) - A_1(j, \omega - \Omega)]}, \quad (5.39)
\]

and \(\bar{x}_{jj}(\omega)\) is the momentum average of the IMA\(^{(0)}\) solution at a reduced frequency:

\[
\bar{x}_{jj}(\omega) = \frac{1}{N} \sum_q G_{IMA\,(0)}(k - q, j, j, \omega - \Omega). \quad (5.40)
\]

The IMA\(^{(1)}\) result is again easy to solve numerically because \(v_1(j, \omega)\) goes to zero quickly away from the surface, and \(j' \leq 5\) cutoffs are sufficient. Thus, we can calculate any \(G_{IMA\,(0)}(k, i, j, \omega)\) at either the IMA\(^{(0)}\) and IMA\(^{(1)}\) level by solving numerically a system of only a few coupled linear equations.

As mentioned above, more complicated surface potentials and surfaces can be treated in this formalism, with the only changes occurring in the expression for \(G_s\). This can be generalized to a broader class of problems, which also include disorder and/or impurities [87]. Of course the invariance to translations is then lost, but this complication can be handled. In the next section we will present results for the simplest case of a (100)-like surface, in the presence of both a zero and non-zero surface potential.
5.3 IMA Results

To understand the effects of the surface and its potential we look at their effect on the spectral weight, defined as [see Eq. (5.15)]:

\[
A(k, i, i, \omega) = -\frac{1}{\pi} \text{Im} \, G(k, i, i, \omega).
\]  

(5.41)

For comparison to the bulk results of Chapter 2 we can use our previous expressions for the spectral weight, but summed over all momenta \(k_\perp\), where \(k_\perp\) is taken to be the direction perpendicular to the surface. Because the MA\((0)\) and MA\((1)\) results for the Holstein model are momentum-independent, this summation is trivial, and we can use the same trick that we used in the calculation of the total DOS in Section 2.3 [see Eq. (2.60)]:

\[
A_{\text{bulk}}(k, \omega) = \frac{1}{N_\perp} \sum_{k_\perp} A_{\text{MA}(n)}(k, k_\perp, \omega) = g_{0D}^1(\omega - \epsilon_\parallel k - \Sigma_{\text{MA}(n)}(\omega)),
\]  

(5.42)

where \(k\) is the in-plane momentum of the electron, and \(g_{0D}^1(\omega)\) is the momentum average of the bulk free propagator in 1D, given explicitly in Eq. (2.47). In 1D Eq. (5.42) is the total DOS of Eq. (2.60).

We will present results for \(d = 1\), in which case there is no in-plane momentum-dependence, and for \(d = 2\), where the in-plane momentum is just a scalar \(k\). We are in the process of obtaining 3D results, but as we show below, the behaviour in 1D and 2D is universal and we expect the 3D results to be very similar to those presented here. We begin by considering the case of isotropic hopping: \(t = t_\parallel\).

5.3.1 Isotropic Hopping

In Fig. 5.2 we look at the 1D spectral weight of the Holstein polaron near the chain’s end in the absence of a surface potential, with a phonon energy \(\Omega/t = 0.5\) and an intermediate coupling strength of \(\lambda = 1\). The top 4 panels show the spectral weight \(A(i, i, \omega)\) as a function of \(\omega\) for the first 4 sites. The bottom panel shows the convergence towards the bulk result. Both the IMA\((0)\) (dashed red line) and IMA\((1)\) (solid blue line) results are shown, in good agreement with each other, but very different from the bulk result. In the figure we see that for the first few layers there is very little weight in the polaron band, especially for \(i = 0\) and \(i = 1\). In the IMA\((1)\) case, this missing spectral weight appears as three distinct surface states near \(\omega/t \approx -2.3, -1.8, \text{ and } -1.3\). These surface states will be discussed.
5.3. IMA Results

Figure 5.2: $A(i, i, \omega)$ vs. $\omega/t$ for a Holstein polaron near a surface (chain’s end) in the absence of a surface potential, shown for various layers starting at the surface $i = 0$. Results are shown using IMA$(0)$ and IMA$(1)$ (shifted for clarity), and compared to the MA$(1)$ bulk result in the bottom panel. All the plots are shown on the same scale. The parameters are $\Omega = 0.5t$, $\lambda = 1$, $U = 0$, and $\eta = 0.004t$.

at length later on. As $i$ increases the $A(i, i, \omega)$ starts to recover towards its bulk value, although full convergence is not reached even for $i = 9$.

We point out that the bulk result shown in the figure is actually the total DOS from Eq. (5.42), and therefore the spectral features have large widths from the integration of $A_{MA(0)}(k_\perp, \omega)$ over $k_\perp$. For comparison, the bulk 1D result $A_{MA(0)}(k_\perp, \omega)$ was shown in Fig. 2.19 for the same set of parameters. In Fig. 5.2 the first polaron band begins at the lowest bound state, at $\omega/t \approx -2.75$, and extends to $\omega/t \approx -2.5$. The second bound state at $\omega/t \approx -2.3$ signifies the beginning of the second polaron band, extending to $\omega/t \approx -2$, and so on. Because this is the MA$(1)$ result, we know that we must produce the polaron+one-phonon continuum beginning at $E_{GS} + \Omega$ as well. However, because the continuum has a fairly small weight it cannot be distinguished from the higher weight features of the second polaron band.
5.3. IMA Results

Figure 5.3: \( A(k=0, i, i, \omega) \) vs. \( \omega/t \) for a Holstein polaron near a surface in the absence of a surface potential, shown for various layers starting at the surface \( i = 0 \). Results are for IMA\(^{(0)}\) and IMA\(^{(1)}\) (shifted for clarity) and compared to the MA\(^{(1)}\) bulk result. All the plots are shown on the same scale. The parameters are \( \Omega = 0.5t \), \( \lambda = 1 \), \( U = 0 \), \( \eta = 0.004t \), and \( t = t_{||} \). The results in the bottom panel are obtained using IMA\(^{(1)}\).

for the parameters in this plot.

Similar plots can be obtained in 2D. In Fig. 5.3 we again consider the case with no surface potential and an intermediate coupling \( \lambda = 1 \). In the top 5 panels of Fig. 5.3 we show the spectral weight \( A(k = 0, i, i, \omega) \) as a function of \( \omega \), for the first 5 layers. The bottom panel shows the convergence towards the bulk result for large \( i \). As in the 1D case, we see that for the first few layers there is very little weight in the polaron band. In this case the weight missing from the polaron band for \( i = 0 \) appears as two bound states near \( \omega/t \approx -4 \) and \(-3.5\), whose weight decreases exponentially as \( i \) increases. These are again surface states, located where small gaps appear in the bulk spectrum, and are the most visible features for the \( i = 0 \) and
5.3. IMA Results

\( i = 1 \) layers. The bulk result of Fig. 5.3 can also be compared to Fig. 2.22, where we plot the spectral function \( A^{(n)}(k, k_{\perp}, \omega) \) for the same set of parameters.

We can explicitly calculate the weight in these surface states by numerically integrating the spectral weights shown Fig. 5.3. The entire spectral weight is unity, as expected because IMA\(^{(n)}\) satisfies the first six sum rules exactly (it is the zeroth order sum rule that guarantees that \( A(k, i, i, \omega) \) is normalized, see Section 2.2.3 for details). The weights in the IMA\(^{(1)}\) surface states at \( \omega/t \approx -4 \) and \( -3.5 \) for \( i = 0 \) are approximately 0.026 and 0.022, respectively, while the weight of the entire first polaron band is only 0.016. This is significantly different from the bulk case, where the first polaron band is the most dominant feature, with a weight of 0.092.

We also find that the location of these surface states depends on the surface potential \( U \). In Fig. 5.4 we show that for a small \( U/t = 0.5 \), surface states appear below the polaron band and in the first gap. Apart from the surface states, which are the most visible features, the remaining spectral weights for \( U = 0 \) and \( U/t = 0.5 \) are fairly similar, the latter with somewhat more weight at lower energies (the vertical scale is the same for all panels). For \( i \gg 1 \) both converge to the same bulk value, plotted in both Figs. 5.3 and 5.4.

The appearance of these surface states is again due to the additional surface potential \( v_n(i, \omega) \) in Eqs. (5.35) and (5.37). This is demonstrated in Fig. 5.5, where we plot the full IMA\(^{(1)}\) solution (solid blue line) and the solution corresponding to \( v_1(i, \omega) = 0 \) (dashed orange line) for the first two layers. In the latter case there are no discrete surface states, for either \( U = 0 \) or \( U/t = 0.5 \), showing that the surface states arise because of the effective potential \( v_n(i, \omega) \). This is in agreement with the fact that \( U < t \) alone cannot bind a surface state. In fact, the relevant ratio is \( U/t^* \), where \( t^* \) is linked to the effective polaron mass [see Eq. (1.34)]. In Figs. 5.4 and 5.5(b) this ratio is still less than 1. The same behaviour is found in \( d = 1 \) and for small/large couplings.

One can also track the lowest bound state as a function of \( U \) to determine the minimum value of \( U \) at which a bound state appears below the first polaron band. This is shown in Fig. 5.6, along with the associated quasiparticle weight in this bound state. We find that the minimum value of \( U \) required to form a bound state decreases with increasing el-ph coupling, as one would expect. Once formed, the energy of this lowest bound state decreases linearly with increasing \( U \). The quasiparticle weight of the lowest bound state is found to reach a maximum just above the minimum value of \( U \), and then to decrease with increasing \( U \) as the spectral weight is shifted.
5.3. IMA Results

Figure 5.4: The 2D spectral weight $A(k=0, i, i, \omega)$ vs. $\omega/t$ for a Holstein polaron near a surface with a surface potential, shown for various layers starting at the surface $i = 0$. Results are for IMA$^{(0)}$ and IMA$^{(1)}$ (shifted for clarity), and compared to the MA$^{(1)}$ bulk result. All plots are shown on the same scale. The parameters are $\Omega = 0.5t$, $\lambda = 1$, $U/t = 0.5$, $\eta = 0.004t$, and $t = t_{\|}$. The results in the bottom panel are obtained using IMA$^{(1)}$.

to additional bound states that appear with a larger surface potential.

To better understand the origin of the additional surface states produced by the effective potential $v_{n}(i, \omega)$ we explicitly plot its components. This is done in Fig. 5.7. In the top two panels we plot the real and imaginary parts of $v_{1}(0, \omega) = \Sigma_{IMA^{(1)}}(0, \omega) - \Sigma_{MA^{(1)}}(\omega)$. Only results for the first layer are shown, but for larger $i$ we find that $v_{n}(i, \omega)$ decays to zero very rapidly, as already discussed. From Fig. 5.7 we see that the bulk self-energy can be quite large, since the bulk polaron spectrum is renormalized considerably compared to the free electron. The “surface” self-energy is similarly large, but its main features are displaced because the Green’s function near the surface is smaller if $U = 0$. When the two are subtracted to obtain $v_{1}$, we
5.3. IMA Results

Find that the real part of the effective potential alternates between positive and negative values. To bind a surface state, this potential has to be negative where the bulk spectrum has a gap. When these two conditions are met, a surface state is likely to appear. Also, in the bottom panel of Fig. 5.7 we plot the surface spectral weight $A(k=0,0,0,\omega)$ and bulk spectral weight $A_{\text{bulk}}(k=0,\omega)$, showing the correlation between the self-energies and the spectral weight.

Using our efficient and accurate approximation, we can generate theoretical predictions such as those in Fig. 5.8, which shows the dispersion of the spectral weight as a function of the in-plane momentum $k$, for the first 3 layers. These plots are obtained using IMA(1) with $U/t = 0.5$, and because the intensity scales are the same on all plots this figure clearly shows that the spectral weight is dominated by surface states for $i = 0$ and $i = 1$. We also observe that the weight of these surface states drops off exponentially for all momenta and that their dispersion is fairly flat.

Such data can be generated very quickly using IMA and compared against experiment to extract accurate values for parameters, which nei-
5.3. IMA Results

Figure 5.6: (a) The energy of the lowest bound state below the first polaron band and (b) its associated quasiparticle weight as a function of $U/t$. The results are for the surface layer in 2D and calculated using IMA\(^{(0)}\) with $\Omega/t = 0.5$ and $t = t_\parallel$.

ther fitting to bulk values nor naive models of quasiparticles near surfaces can hope to capture. If one did not expect surface states to appear and interpreted these features as being bulk polarons, one might extract wrong parameters from fitting such data.

5.3.2 Anisotropic Hopping

As mentioned in Section 5.1, it is straightforward to study the effects of anisotropic hopping because of the clear distinction between the in-plane and inter-layer hopping integrals appearing in Eq. (5.14). In the limit of $t/t_\parallel \to 0$ we expect each layer to become fully decoupled, independently reproducing the bulk planar result in each layer. This is shown in Fig. 5.9 for the surface layer. The plots are for 2D and an intermediate coupling strength of $\lambda = 1$. The spectral weight of the surface layer ($i = 0$) is plotted for $t/t_\parallel = 1, 0.95, 0.75, 0.25$ and 0.1. We find that as the inter-layer hopping is decreased the poles shift toward the 1D bulk result. For $t = 0.25t_\parallel$ the spectral weight is almost completely localized around the 1D bulk result pole.
5.3. IMA Results

Figure 5.7: The components of the effective potential $v_1(i = 0, \omega)$ vs. $\omega$ in the absence of a surface potential. The parameters are $\Omega/t = 0.5$, $\lambda = 1$, $U = 0$, $\eta = 0.004t$, and $t = t_{||}$. The top two panels show the real and imaginary parts of the surface and bulk components of $v_1(0, \omega) = \Sigma_{IMA}(1)(0, \omega) - \Sigma_{MA}(1)(\omega)$, and the bottom panel shows a comparison of the bulk and surface ($i = 0$) spectral weights for the same parameters.

locations, and for $t = 0.1t_{||}$ the 2D surface result and the 1D bulk result are almost indistinguishable. Similar results are found for $i > 0$. Note that the 1D bulk result is shown with $\lambda = 2$ to account for the factor of $d$ in the definition $\lambda = g^2/(2d\Omega t)$ [see Eq. (1.46)].

In Fig. 5.10 we plot the 2D spectral weight at the surface for various values of $t/t_{||}$ in the presence of an attractive surface potential of strength $U/t = 0.5$. We again find that the 2D IMA result approaches the bulk 1D result as $t \to 0$, but at a reduced energy $\omega - U$, i.e. in the limit of $t/t_{||} \to 0$ the surface potential simply shifts the energies by $U$ at the surface layer. Since we have modeled the surface potential as being localized on the surface only, in the limit of $t \to 0$ only the energies of the surface layer are shifted, however, for a more long range surface potential and/or finite $t$ the
5.3. IMA Results

Figure 5.8: The spectral weight $A(k, i, i, \omega)$ vs. the in-plane momentum $k$ and energy $\omega$, for the first 3 surface layers and bulk. The results are obtained from IMA(1), and the parameters are $\Omega/t = 0.5t$, $\lambda = 1$, $U/t = 0.5$, $\eta = 0.004t$, and $t = t_{||}$. The intensity scales are the same in all plots, and the surface states are the dominant features for $i = 0$ and $i = 1$.

In a manner similar to Fig. 5.6, we can also find the minimum surface potential $U/t_{||}$ for which a bound state forms below the first polaron band, for various values of $t/t_{||}$. This is shown in Fig. 5.11 for an intermediate coupling strength of $\lambda = 1$. For a given ratio of $t/t_{||}$, we find that the energy of this bound state decreases linearly with $U/t$ after some minimum value of $U/t$ required to form a bound state, as found previously for the $t/t_{||}$ case. We also find that this minimum value of $U/t$ decreases as $t/t_{||} \rightarrow 0$. This can be understood as follows. In the limit of $t \rightarrow 0$ the $d$-dimensional surface result reduces to the $(d - 1)$-dimensional bulk result at a reduced frequency: $A(k, \omega - U)$. Since the bottom of the first polaron band is determined from the $U = 0$ result, we are guaranteed a bound state for any finite $U$ in this limit. Therefore, for a small but finite $t/t_{||}$, we require only a very small value of $U/t$ to form a bound state below the first polaron band at the surface layer, as shown in Fig. 5.11. In the right-hand side of Fig. 5.11 we show the quasiparticle weight in this bound state for several values of $t/t_{||}$. The weight in this bound state is found to decrease with decreasing $t$ until
5.4 Conclusion

In this chapter we have shown how to treat the problem of a Holstein polaron near a surface using the inhomogeneous MA approximation. This
5.4. Conclusion

Figure 5.10: The 2D spectral weight $A(k=0, i, i, \omega)$ vs. $\omega$ at the surface for various inter-layer hopping values $t/t_{||}$ and $U/t = 0.5$. The results are obtained from $\text{IMA}^{(0)}$ and the parameters are $\Omega/t = 0.5$, $\lambda = 1$, $\eta = 0.004t$, and $t/t_{||} = 1, 0.95, 0.75, 0.25, 0.1$. The solid black line shows the 1D bulk result $A(k=0, \omega - U)$ with $\lambda = 2$.

approximation allows us to treat the broken symmetry without any further approximations, and possesses the same properties and level of accuracy as the previous MA approximations. The main result that we found was that additional surface states can appear, even in the absence of a surface potential, for both $d = 1$ and $d = 2$. A close inspection of these surface states revealed that the surface potential is strongly renormalized via the el-ph interactions, and it can thus bind surface states even when the bare surface potential cannot do so. These surface states are the dominant spectral features for the first two layers and illustrate that the interpretation of experimentally obtained surface data is not straightforward. We have also applied this approximation to treat anisotropic hopping, which is particu-
5.4. Conclusion

Figure 5.11: (a) The energy of the lowest bound state below the first polaron band and (b) its associated quasiparticle weight as a function of $U/t$. The results are for the surface layer in 2D and calculated using $IMA^{(0)}$ with $\Omega/t = 0.5$ and $t/t = 1, 0.75, 0.5, 0.25$ and 0.1.

larly relevant in layered materials. We have shown that the planar bulk result is recovered in the limit of small inter-layer hopping, as expected, and that the relevant ratio for binding a surface state is related to the ratio of the surface potential strength and the inter-layer hopping. We plan to study the effects of these surface states and anisotropy in more detail in a future work.
Chapter 6

Conclusion

6.1 Summary of this Work

In this work we have presented a highly efficient and accurate approximation that can be used to calculate the Green’s function of a polaron: the momentum average (MA) approximation. This approximation is unique because it sums all of the terms in the full diagrammatical expansion of the self-energy, albeit with approximations made on almost all of them so that this sum can actually be performed. It is exact in both the zero coupling and zero bandwidth limits, and it is found to be accurate everywhere in parameter space. In this thesis we have applied MA to a handful of el-ph coupling models, studied its interpretation and accuracy extensively, and shown how the approximation can be systematically improved and modified to account for broken translational symmetry. We summarize the key results from each chapter below.

In Chapter 2 we developed the MA approximation, first using it to investigate the Holstein model. We derived the exact equations of motion for the Holstein polaron and then showed that by replacing all free propagators by their momentum averages in the self-energy expansion we could obtain an analytical solution for the Green’s function. This was followed by a detailed analysis of the approximation’s physical meaning and accuracy. It was easy to show that MA was exact for both zero coupling \((g = 0)\) and zero bandwidth \((t = 0)\). Through a diagrammatical analysis we showed that MA treats the first diagram exactly, sums all terms in the diagrammatical expansion, and captures most of the \(t\) dependence in all diagrams. A more stringent test of MA followed, via the spectral weight sum rules. We demonstrated that MA satisfies the first six sum rules exactly, and that it remains accurate for all higher sum rules, becoming exact in both asymptotic limits mentioned above. Finally, we presented an alternative interpretation for the meaning of the MA approximation. By considering the real space interpretation of momentum averaging the free propagators appearing in the self-energy, we showed that MA can be thought of as a variational method with basis states having phonons on a single site only. This interpretation...
6.1. Summary of this Work

turned out to be very important in later chapters where we systematically improved and generalized the MA approximation.

In the remainder of Chapter 2 we further gauged the success of the MA approximation through extensive comparisons to existing numerical data. There is a large amount of such data for the Holstein model, including ground state properties, momentum-dependent results, and some spectral (higher-energy) data. The MA approximation was compared to these results and it was shown that the approximation was able to produce highly accurate results over most of parameter space, as expected based on the arguments above. However, these comparisons did reveal a few shortcomings of MA. The accuracy of MA was found to decrease when the phonon energy $\Omega \rightarrow 0$, and MA either predicted the incorrect location for the electron+one-phonon continuum, or no continuum at all. These limitations provided the main motivation for systematically improving the MA approximation in the following chapter.

In Chapter 3 we addressed these shortcomings by developing a series of improved approximations which we called MA$^{(0)}$, MA$^{(1)}$, MA$^{(2)}$ and so on (the original MA approximation is MA$^{(0)}$ in this notation). This was done by recognizing that in real space the MA approximation keeps only the diagonal contributions of the free propagator $G_0(i, j, \omega - n\Omega)$ for $n > 0$. The ignored contributions decay with the distance $|i - j|$, at least for low energies, and this decay is steeper for larger $n$. Therefore, higher order systematic improvements MA$^{(n)}$ were obtained by keeping the propagators with energy $\omega - m\Omega$ exactly in all self-energy diagrams, where $m \leq n$.

At the MA$^{(1)}$ level of approximation we were able to produce the correct location for the electron+one-phonon continuum. This was understood in variational terms, as it was shown that MA$^{(1)}$ roughly corresponds to a variational basis where one phonon is allowed to be away from the main polaron cloud, thus always producing the correct location for the electron+one-phonon continuum. At the MA$^{(2)}$ level we obtained a weakly momentum-dependent self-energy. This was missing from the first two levels of approximation because of the featureless nature of the Holstein model (both the phonon energy and el-ph coupling are momentum-independent in the Holstein model, for other models a momentum-dependent self-energy is found even at the MA$^{(0)}$ level). In addition to remedying the problem with the electron+one-phonon continuum, the higher order MA$^{(n)}$ approximations gave improved accuracy everywhere in parameter space and satisfied even higher order sum rules exactly. The most notable improvements were found for lower dimensions and small $\Omega$. Therefore, all shortcomings of the MA$^{(0)}$ approximation were remedied by systematically improving the approxima-
6.1. Summary of this Work

In Chapter 4 we generalized the applicability of the MA approximation to more general el-ph coupling models. We first derived a simple generalization applicable to models possessing a general momentum-dependent el-ph coupling of the form $g(q)$. While this simple generalization was extremely efficient and fairly accurate when applied to the available numerical data for the 1D breathing-mode Hamiltonian, it was non-variational and failed to reproduce some of the features captured by the numerical simulations. Through variational considerations it was then shown that we could restore the variational nature of the MA approximation and considerably improve its accuracy. Using the 1D breathing-mode Hamiltonian as an explicit example, we showed that by using this method we were able to obtain ground state energies with at most 0.3% error from the available numerical data. Also, by systematically improving both the simple and variational generalizations of MA, we were able to increase the accuracy of the approximations and produce the correct location for the electron+one-phonon continuum, as in the Holstein case. In the last half of the chapter we investigated several properties for the 1D breathing-mode Hamiltonian using the variational MA approximation. These included both ground state and higher energy properties, and comparisons were made to numerical data when possible.

In Chapter 5 we looked at the effect of a surface on the spectra of a Holstein polaron using another generalization of the MA approximation: the inhomogeneous MA (IMA) approximation. We showed that the broken translational symmetry could be accounted for exactly in our simple surface model involving a (100)-like surface and an on-site constant surface potential, and argued that more complicated surfaces and impurities could also be treated exactly within our formalism. The effect of the el-ph coupling was approximated in an analogous manner to the previous applications of MA, and the resulting IMA approximation is accurate for all coupling strengths. We also showed that the IMA approximation could be systematically improved, as before, and we explicitly derived the IMA$^{(1)}$ expression.

The main results of Chapter 5 were that the surface changed the spectral properties significantly, with bulk properties being recovered only very far away from it, and that the renormalization of the surface potential by the el-ph coupling gave rise to an additional surface potential. This additional surface potential was found to be responsible for binding additional surface states where they are not normally expected, and was present for all dimensions and parameters studied. These results demonstrated that the interpretation of data obtained from surface spectroscopies that probe only the first few layers of a material is not necessarily simple or straightforward.
Since the original proposal of the MA approximation almost four years ago [90] its accuracy has been systematically improved and it has been generalized to a wide variety of systems. In addition to the work discussed in this thesis, MA has been used to study multiple phonon modes [15], graphene [91], phonon statistics [92], and systems with Rashba spin-orbit coupling [93]. More recently, the IMA approximation has been developed to treat broken translational symmetry. This has been done in the presence of disorder [87], and in the case of surfaces, as discussed in Chapter 5.

There are several possible applications of MA that have yet to be investigated. The first class of problems would be to consider more complicated materials described by models of the same type as presented here. In these cases we could apply our existing formalism in a straightforward fashion to describe more realistic systems. Examples could include materials involving momentum-dependent el-ph coupling, disorder, surfaces, interfaces, multiple phonon modes, or any combination of the above. An interesting problem would be to apply the IMA approximation to more complicated surfaces and to allow for the presence of impurities on the surface (i.e. adsorbed atoms). Such problems are extremely difficult to treat numerically, but since the broken translational invariance can be treated exactly within IMA, we would expect very accurate results using our approximation.

There is also a second class of problems involving further generalizations of the MA approximation. Examples would include applying MA to calculate optical conductivity and other response functions, generalizing to finite temperature, treating phonon dispersion, allowing for phonon-modulated hopping, and treating many-polaron systems, i.e. finite electron concentrations. Of course this is a huge class of problems, and we do not know which of these would actually permit a MA-like approximation, but it would be very interesting to see which ones could be treated using MA.

Given its rapidly increasing range of applicability over the last few years, MA is already a useful tool for treating problems that are difficult, or even impossible to solve using numerical techniques. Given its accurate treatment of surfaces and disorder, it will also be of interest to experimentalists looking to interpret data from systems where polaronic effects are important. It is our hope that MA will continue to be a useful analytical tool for many years to come.
Bibliography


Appendix A

Comparison with DMFT

The MA self-energy of Eq. (2.25) looks similar to the DMFT self-energy, discussed in Ref. [71]. This is not so surprising, since both become equal to the exact Lang-Firsov limit if the bandwidth goes to zero, and this can be rewritten as an infinite continued fraction, as shown in Eq. (2.17). However, this similarity may raise questions about the relationship between the two approximations. In this appendix, we briefly show that the two approximations are very different at all finite $t$ and $g$.

The meaning of the MA $\tilde{g}_0(\omega)$ and of corresponding DMFT $G_0(\omega)$ (using the notation of Ref. [71]) is very different. The DMFT $G_0(\omega)$ is obtained by solving exactly the problem of an impurity coupled to an environment in the $d \to \infty$ limit, and then imposing the self-consistency condition that the impurity site behaves similarly to all other sites in the environment. The DMFT $G_0(\omega)$ is calculated self-consistently using the following steps [71]: (i) with some initial guess for $G_0(\omega)$, one calculates the DMFT self-energy, given by a formula similar to Eq. (2.25), with $\tilde{g}_0(\omega)$ replaced by $G_0(\omega)$; this self-energy is then used to calculate the total Green’s function

$$G(\omega) = \int_{-\infty}^{\infty} d\varepsilon \rho_0(\varepsilon) \frac{1}{\omega - \varepsilon - \Sigma(\omega) + i\eta}, \quad (A.1)$$

where $\rho_0(\varepsilon)$ is the density of states of the non-interacting electrons. The usual procedure is to take $d \to \infty$ and use as DOS the semielliptical formula corresponding to an infinitely branched Bethe tree,

$$\rho_0(\varepsilon) = \frac{2}{\pi \left(\frac{W}{2}\right)^2} \sqrt{\left(\frac{W}{2}\right)^2 - \varepsilon^2}, \quad (A.2)$$

where $W$ is the bandwidth for the non-interacting system. Then, (iii) the new $G_0(\omega)$ is extracted from the condition that $G(\omega) = \left[G_0^{-1}(\omega) - \Sigma(\omega)\right]^{-1}$ and the procedure is repeated until self-consistency is reached. Reaching self-consistency is a non-trivial numerical task, especially compared to obtaining the MA $\tilde{g}_0(\omega)$ (see below). More importantly, the DMFT $G_0(\omega)$ is an explicit function of $g$ and $\Omega$. 

154
Appendix A. Comparison with DMFT

Figure A.1: Comparison between the function $G_0(\omega)$ entering the DMFT self-energy, for $\omega/t = 0.5$ and $\lambda = 0.5$ (green) and $\lambda = 1.5$ (yellow), and $\bar{g}_0(\omega)$ entering the MA self-energy, for the $d \to \infty$ semi-elliptical DOS.

In contrast, the MA $\bar{g}_0(\omega)$ is the momentum average of the free propagator: thus, it is a known function, independent of the parameters $g$ and $\Omega$. In particular, for the semi-elliptical DOS, we have simply:

$$
\bar{g}_0(\omega) = \int_{-\infty}^{\infty} d\varepsilon \frac{\rho_0(\varepsilon)}{\omega - \varepsilon + i\eta} = \frac{8(\omega + i\eta)}{W^2} \left[ 1 - \sqrt{1 - \frac{W^2}{4(\omega + i\eta)^2}} \right]. \quad (A.3)
$$

A comparison of these functions is provided in Fig. A.1. They are clearly different. Moreover, note that in the MA approximation, $G(k, \omega)$ is an explicit function of the momentum. The MA self-energy for the Holstein polaron happens to be independent of the momentum, but this a consequence of the simplicity of the model, not an “in-built” feature like in DMFT, and by systematically improving MA one does obtain an explicitly momentum-dependent self-energy for the Holstein model. Generalizations to models with a momentum-dependent coupling and/or dispersive phonons lead to momentum-dependent self-energies as well.
Appendix B

MA(1) Details

Our goal is to solve for $\Sigma_{MA(1)}$, given by the recurrence relations in Eqs. (3.4)-(3.6).

We introduce the partial momentum averages:

$$\bar{f}_n(q_1) = \frac{1}{N-n-1} \sum_{q_2,\ldots,q_n} f_n^{(1)}(q_1,\ldots,q_n), \quad (B.1)$$

where we use the convention that $\bar{f}_1(q_1) = f_1^{(1)}(q_1)$. We also relate this to our previous definition of the full momentum averages [see Eq. (2.19)]:

$$\mathcal{F}_n = \frac{1}{N^n} \sum_{q_1,\ldots,q_n} f_n^{(1)}(q_1,\ldots,q_n) = \frac{1}{N} \sum_{q_1} \bar{f}_n(q_1), \quad (B.2)$$

so that $\Sigma_{MA(1)}(k,\omega) = \mathcal{F}_1$.

First, we average each of Eqs. (3.6) over all the corresponding momenta, to find that for $n \geq 2$,

$$\bar{f}_n(q_1) = \tilde{g}_0(\omega - n\Omega)[ng^2\mathcal{F}_{n-1} + \mathcal{F}_{n+1}].$$

This has the following continued fraction solution [see Eq. (2.23)]:

$$\mathcal{F}_2 = g^2A_2(\omega)\mathcal{F}_1, \quad (B.3)$$

where $A_2(\omega)$ is a continued fraction defined in Eq. (2.26) and $\mathcal{F}_1$ is still unknown. We now average each of Eqs. (3.6) over all the momenta except $q_1$, to find that for all $n \geq 2$

$$\bar{f}_n(q_1) = \tilde{g}_0(\omega - n\Omega) \left[g^2\mathcal{F}_{n-1} + (n-1)g^2\bar{f}_{n-1}(q_1) + \bar{f}_{n+1}(q_1)\right]. \quad (B.4)$$

This can also be solved by rewriting it in terms of

$$\delta \bar{f}_n(q_1) = \bar{f}_n(q_1) - \mathcal{F}_n. \quad (B.5)$$

Using the recurrence relation for $\mathcal{F}_n$, we find that:

$$\delta \bar{f}_n(q_1) = \tilde{g}_0(\omega - n\Omega) \left[(n-1)g^2\delta \bar{f}_{n-1}(q_1) + \delta \bar{f}_{n+1}(q_1)\right], \quad (B.6)$$
which also has solutions in terms of continued fractions:

$$\delta f_2(q_1) = g^2 A_1(\omega - \Omega) \delta f_1(q_1). \quad (B.7)$$

Using Eq. (B.3) we then find

$$\bar{f}_2(q_1) = g^2 A_1(\omega - \Omega) f_1(q_1) + g^2 [A_2(\omega) - A_1(\omega - \Omega)] f_1. \quad (B.8)$$

This expression is now inserted in Eq. (3.5) to give an equation with only \( f_1^{(1)}(q_1) \) and \( F_1 = \frac{1}{N} \sum_{q_1} f_1^{(1)}(q_1) \) as unknowns, which can be re-arranged as:

$$\left( [G_0(k - q_1, \omega - \Omega)]^{-1} - g^2 A_1(\omega - \Omega) \right) f_1^{(1)}(q_1)
= g^2 + g^2 [A_2(\omega) - A_1(\omega - \Omega)] F_1. \quad (B.9)$$

By then using the following identity

$$[G_0(k, \omega)]^{-1} - a(\omega) = \omega - \varepsilon_k + i\eta - a(\omega) = [G_0(k, \omega - a(\omega))]^{-1}, \quad (B.10)$$

we find:

$$f_1^{(1)}(q_1) = G_0(k - q_1, \omega - \Omega - g^2 A_1(\omega - \Omega))
\times [g^2 + g^2 [A_2(\omega) - A_1(\omega - \Omega)] F_1]. \quad (B.11)$$

Finally, after momentum averaging over \( q_1 \) on both sides, we find obtain the desired MA\(^{(1)}\) result for \( \Sigma_{MA^{(1)}} = F_1 = \frac{1}{N} \sum_{q_1} f_1^{(1)}(q_1) \), given in Eq. 3.7.
Appendix C

MA\(^{(2)}\) Details

As for MA\(^{(1)}\), the strategy is to solve Eqs. (3.14) to find how \(\frac{1}{N} \sum_{q_3} f_3^{(2)}(q_1, q_2, q_3)\) depends on \(f_2^{(2)}(q_1, q_2)\). Using this in Eq. (3.13) allows us to solve it in conjunction with Eq. (3.12) to find \(f_1^{(2)}(q_1)\) and therefore \(\Sigma\).

We introduce various partial momentum averages:

\[
\bar{f}_n(q_1, q_2) = \frac{1}{N^{n-1}} \sum_{q_2, \ldots, q_n} f_n^{(2)}(q_1, q_2, \ldots, q_n),
\]

\[
\bar{f}_n(q_1) = \frac{1}{N^{n-1}} \sum_{q_2, \ldots, q_n} f_n^{(2)}(q_1, \ldots, q_n),
\]

where we take \(\bar{f}_1(q_1) = \bar{f}_1^{(2)}(q_1)\), and \(\bar{f}_2(q_1, q_2) = f_2^{(2)}(q_1, q_2)\). We again recall the definition of the full momentum average:

\[
F_n = \frac{1}{N} \sum_{q_1, \ldots, q_n} f_n^{(2)}(q_1, \ldots, q_n) = \frac{1}{N} \sum_{q_1} \bar{f}_n(q_1).
\]

The functions \(\bar{f}_n, \bar{f}_n\) and \(F_n\) should also carry the upper label (2) since their values are different than those obtained within MA\(^{(1)}\), however we do not write it explicitly to simplify the notation somewhat.

The solutions for \(F_n\) and \(\bar{f}_n\) proceed just as in the MA\(^{(1)}\) case, with one difference. Averaging all equations (3.14) over all corresponding momenta, we find that now only for \(n \geq 3\), \(F_n = g_0(\omega - n\Omega)[ng^2F_{n-1} + F_{n+1}]\), therefore this recurrence relation now ends with:

\[
F_3 = g^2A_3(\omega)F_2.
\]

Similarly, after using the same approach as for MA\(^{(1)}\), one finds that

\[
\delta \bar{f}_3(q_1) = \bar{f}_3(q_1) - F_3 = g^2A_2(\omega - \Omega)\delta \bar{f}_2(q_1)
\]

so that

\[
\bar{f}_3(q_1) = g^2A_2(\omega - \Omega)\bar{f}_2(q_1) + g^2[A_3(\omega) - A_2(\omega - \Omega)]F_2.
\]
Appendix C. MA(2) Details

Of course, the functions $\mathcal{F}_2, \bar{f}_2(q_1)$ are still unknown. Finally, we can proceed to calculate $\bar{f}_3(q_1, q_2)$ which is needed in Eq. (3.13).

We momentum average Eqs. (3.14) over all momenta except $q_1, q_2$, to find

$$\bar{f}_n(q_1, q_2) = \bar{g}_0(\omega - n\Omega) \left( g^2 [f_{n-1}(q_1) + \bar{f}_{n-1}(q_2)] + (n - 2)g^2 \bar{f}_{n-1}(q_1, q_2) + \bar{f}_{n+1}(q_1, q_2) \right). \quad (C.7)$$

Now define

$$\delta \bar{f}_n(q_1, q_2) = \bar{f}_n(q_1, q_2) - \bar{f}_n(q_1) - \bar{f}_n(q_2) + \mathcal{F}_n. \quad (C.8)$$

For these, the recurrence relations are

$$\delta \bar{f}_n(q_1, q_2) = \bar{g}_0(\omega - n\Omega) \left[ (n - 2)g^2 \delta \bar{f}_{n-1}(q_1, q_2) + \delta \bar{f}_{n+1}(q_1, q_2) \right] \quad (C.9)$$

with the solution

$$\delta \bar{f}_3(q_1, q_2) = g^2 A_1(\omega - 2\Omega) \delta \bar{f}_2(q_1, q_2), \quad (C.10)$$

from which we find

$$\bar{f}_3(q_1, q_2) = g^2 A_1 f^{(2)}_2(q_1, q_2) + g^2 (A_2 - A_1) [\delta \bar{f}_2(q_1) + \delta \bar{f}_2(q_2)] + g^2 [A_3 - A_1] \mathcal{F}_2, \quad (C.11)$$

i.e. only in terms of various partial averages of $f^{(2)}_2(q_1, q_2)$. Throughout we use the shorthand notation

$$A_1 \equiv A_1(\omega - 2\Omega), \quad A_2 \equiv A_2(\omega - \Omega), \quad A_3 \equiv A_3(\omega). \quad (C.12)$$

We now substitute $\bar{f}_3(q_1, q_2)$ in Eq. (3.13), which gives:

$$f_2(q_1, q_2) = g^2 G_0(k - q_1 - q_2, \tilde{\omega}) \left( f^{(2)}_1(q_1) + f^{(2)}_1(q_2) \right) + (A_2 - A_1) [\delta \bar{f}_2(q_1) + \delta \bar{f}_2(q_2)] + (A_3 - A_1) \mathcal{F}_2. \quad (C.13)$$

We also used the fact that

$$[G_0(k - q_1 - q_2, \omega - 2\Omega)]^{-1} - g^2 A_1 = [G_0(k - q_1 - q_2, \tilde{\omega})]^{-1}. \quad (C.14)$$
Appendix C. MA$^{(2)}$ Details

where $\tilde{\omega} = \omega - 2\Omega - g^2 A_1$.

Momentum averaging Eq. (C.13) over both momenta leads immediately to $F_2 = g^2 \bar{g}_0(\tilde{\omega}) [2F_1 + (A_3 - A_1)F_2]$, and therefore:

$$F_2 = \frac{2g^2 \bar{g}_0(\tilde{\omega})F_1}{1 - g^2 \bar{g}_0(\tilde{\omega})[A_3 - A_1]},$$

(C.15)

It follows that all higher order total averages $F_n$ are proportional to $F_1 = \Sigma$.

Since only $\bar{f}_2(q_1)$ is needed in Eq. (3.12), we can obtain it by momentum averaging Eq. (C.13) over $q_2$. This leads directly to the closed system of coupled equations:

$$f_1^{(2)}(q_1) = G_0(k - q_1, \omega - \Omega) \left[ g^2 + \delta \bar{f}_2(q_1) + F_2 \right],$$

(C.16)

$$\delta \bar{f}_2(q_1) = g^2 \bar{g}_0(\tilde{\omega}) \left[ f_1^{(2)}(q_1) + (A_2 - A_1)\delta \bar{f}_2(q_1) - 2F_1 \right]$$
$$+ \frac{g^2}{N} \sum_{q_2} G_0(k - q_1 - q_2, \tilde{\omega}) \left[ f_1^{(2)}(q_2) + (A_2 - A_1)\delta \bar{f}_2(q_2) \right].$$

(C.17)

To solve it, we first rewrite these as a single equation in terms of the unknowns $x_q = f_1^{(2)}(q) + (A_2 - A_1)\delta \bar{f}_2(q)$. This is obtained by adding the second equation to $[A_2 - A_1 + G_0(k - q_1, \omega - \Omega)]$ times the first equation. We introduce the short-hand notation:

$$a_{ij} = a_{ij}(\omega) = 1 - g^2 \bar{g}_0(\tilde{\omega})[A_i - A_j]$$

(C.18)

and use the identities

$$[G_0(k - q_1, \omega - \Omega)]^{-1} - g^2 \bar{g}_0(\tilde{\omega}) \left[ 1 + (A_2 - A_1) \right] [G_0(k - q_1, \omega - \Omega)]^{-1}$$
$$= [1 - g^2 \bar{g}_0(\tilde{\omega})(A_2 - A_1)] \left[ G_0(k - q_1, \tilde{\omega}) \right]^{-1},$$

(C.19)

where

$$\tilde{\omega} = \omega - \Omega - \frac{g^2 \bar{g}_0(\tilde{\omega})}{a_{21}},$$

(C.20)

and

$$G_0(k - q_1, \tilde{\omega}) \left[ 1 + (A_2 - A_1) \right] [G_0(k - q_1, \omega - \Omega)]^{-1}$$
$$= A_2 - A_1 + \frac{1}{a_{21}}G_0(k - q_1, \tilde{\omega}).$$

(C.21)
Appendix C. MA\textsuperscript{(2)} Details

Both identities are based on the definition \([G_0(k, \omega)]^{-1} = \omega - \varepsilon_k + i \eta\). With these, the final equation for \(x_\mathbf{q}\) becomes:

\[
\begin{align*}
a_{21} x_\mathbf{q} &= G_0(k - \mathbf{q}, 0) \left[ g^2 + \frac{a_{21} - a_{31}}{a_{21}} F_2 \right] \\
&- a_{31} (A_2 - A_1) F_2 + \left[ A_2 - A_1 + \frac{G_0(k - \mathbf{q}, 0)}{a_{21}} \right] \\
&\times \frac{g^2}{N} \sum_{\mathbf{q}_2} G_0(k - \mathbf{q} - \mathbf{q}_2, 0) x_{\mathbf{q}_2}.
\end{align*}
\]

(C.22)

We now Fourier transform \(x(i) = \frac{1}{N} \sum_{\mathbf{q}} e^{i \mathbf{q} \cdot \mathbf{R}_i} x_{\mathbf{q}}\). From the definition of \(x_\mathbf{q}\) it follows immediately that \(x(0) = F_1 = \Sigma\). Since \(F_2 \sim F_1 = x(0)\) [see Eq. (C.15)], the resulting system of inhomogeneous equations is:

\[
\sum_{j} M_{ij}(k, \omega) x(j) = e^{i k \cdot \mathbf{R}_i} g^2 G_0(-i, 0)
\]

(C.23)

where \(G_0(i, \omega) = \frac{1}{N} \sum_{\mathbf{k}} e^{i \mathbf{k} \cdot \mathbf{R}_i} G_0(\mathbf{k}, \omega)\), and the matrix elements are:

\[
M_{00} = 1 - g^2 \bar{g}_0(0) \bar{g}_0(0) \left( \frac{2}{a_{31}} - \frac{1}{a_{21}} \right),
\]

(C.24)

\[
M_{i0} = -g^2 \bar{g}_0(0) e^{i k \cdot \mathbf{R}_i} G_0(-i, 0) \left( \frac{2}{a_{31}} - \frac{1}{a_{21}} \right)
\]

(C.25)

for \(i \neq 0\), and for both \(i, j \neq 0\):

\[
M_{ij} = a_{21} \delta_{i,j} - g^2 e^{i k \cdot \mathbf{R}_i} G_0(j, 0) \left[ (A_2 - A_1) \delta_{i,-j} + \frac{G_0(-i - j, 0)}{a_{21}} \right].
\]

(C.26)

We truncate this system by keeping only sites \(\mathbf{R}_i\) within a certain distance from the origin (see discussion in main text) and solve it numerically. The self-energy is finally

\[
\Sigma_{\text{MA}(2)}(k, \omega) = x(0).
\]

(C.27)
Appendix D

Momentum Averages

In this appendix we derive the various momentum averages required to evaluate the self-energy expressions for the models with electron-phonon coupling dependent on the phonon momentum. We require the following weighted momentum averages of the free electron propagator for a 1D tight-binding dispersion:

\[ \tilde{g}_j(\omega) \equiv \frac{1}{N} \sum_q e^{\pm i k(a)} G_0(q, \omega). \]  

(D.1)

It is straightforward to show that

\[ \bar{g}_0(\omega) = \frac{\text{sgn}(\omega + i \eta)}{\sqrt{(\omega + i \eta)^2 - 4t^2}}, \]  

(D.2)

\[ \bar{g}_1(\omega) = \frac{1}{2t} \left[ 1 - (\omega + i \eta)\bar{g}_0(\omega) \right], \]  

(D.3)

\[ \bar{g}_2(\omega) = -\bar{g}_0(\omega) - \frac{\omega + i \eta}{t} \bar{g}_1(\omega), \]  

(D.4)

and

\[ \bar{g}_3(\omega) = \frac{1}{t} \left[ 3 - \frac{(\omega + i \eta)^2}{t^2} \right] \bar{g}_1(\omega). \]  

(D.5)

For the MA\(^{(0)}\) and MA\(^{(1)}\) self-energy expressions in Eqs. (4.15) and (4.22) we also require \( \bar{g}_0(k, \omega) \), as defined in Eq. (4.13). For the 1D breathing-mode model one can show that

\[ \bar{g}_0(k, \omega) = 2g^2 \bar{g}_0(\omega) - \frac{g^2 \cos k}{t} \left[ 1 - (\omega + i \eta)\bar{g}_0(\omega) \right]. \]  

(D.6)

In the MA\(^{(v,1)}\) calculation we require the momentum average of Eq. (E.11). For the 1D breathing-mode Hamiltonian [i.e. \( g_q = -2i g \sin(q/2) \) and \( \epsilon_k = -2t \cos(k) \)] we require the following momentum averages:

\[ \frac{1}{N} \sum_q G_0(k - q, \omega) \frac{e^{\pm ni(k-q)}(1 - e^{\mp iq})}{1 - e^{\mp iq}} = -e^{\mp ik} \bar{g}_{n+1}(\omega), \]  

(D.7)
Appendix D. Momentum Averages

\[ \frac{1}{N} \sum_q G_0(k - q, \tilde{\omega}) G_0(k - q, \omega) \frac{e^{\pm n i (k - q)} |q|}{1 - e^{\pm i q}} \]

\[ = \frac{g^2}{\omega - \tilde{\omega}} \left\{ \tilde{g}_n(\tilde{\omega}) - \tilde{g}_n(\omega) - e^{\mp i k} [\tilde{g}_{n+1}(\tilde{\omega}) - \tilde{g}_{n+1}(\omega)] \right\} , \quad \text{(D.8)} \]

and

\[ \frac{1}{N} \sum_q G_0(k - q, \omega) \frac{e^{\pm n i (k - q)} (1 - \cos q)}{1 - e^{\pm i q}} = \frac{1}{2} \left[ \tilde{g}_n(\omega) - e^{\mp i k} \tilde{g}_{n+1}(\omega) \right] . \quad \text{(D.9)} \]
Appendix E

MA\((v, 1)\) Details

Our goal is to find the MA\((v, 1)\) self-energy by solving the recurrence relations of Eqs. (4.19) and (4.42) within the variational basis described by Eqs. (4.27), (4.28), (4.40), and (4.41).

We can immediately use the MA\((v, 0)\) result to solve Eq. (4.19), but only up to the \(n = 2\) level: \(v_2 = Q_2 v_1\). To solve the \(n = 1\) equation exactly we will need to construct a set of recurrence relations involving Eqs. (4.40) and (4.41) using Dyson’s identity. The resulting equations take a form similar to Eqs. (4.31)-(4.36) if we define

\[
\delta f_{n,m}(q_1) = \tilde{f}_{n,m}(q_1) - g(1 - e^{iq_1})f_{n,m} + g(1 - e^{-iq_1})f_{n,m+1}, \quad (E.1)
\]

\[
\delta f_{n,0}(q_1) = \tilde{f}_{n,0}(q_1) - g(1 - e^{iq_1})f_{n,0} + g(1 - e^{-iq_1})f_{n,1} - ge^{-i(k-q_1)(1 - e^{iq_1})}f_{n,n-1}, \quad (E.2)
\]

\[
\delta f_{n,n-1}(q_1) = \tilde{f}_{n,n-1}(q_1) - g(1 - e^{iq_1})f_{n,n-1} + g(1 - e^{-iq_1})f_{n,n} + ge^{i(k-q_1)(1 - e^{-iq_1})}f_{n,n+1}, \quad (E.3)
\]

for \(n > 1\), and

\[
\delta f_{1,0}(q_1) = \tilde{f}_{1,0}(q_1) - g(1 - e^{iq_1})f_{1,0}^{(1)} + g(1 - e^{-iq_1})f_{1,1}^{(1)} - ge^{-i(k-q_1)(1 - e^{iq_1})}f_{1,0}^{(1)} + ge^{i(k-q_1)(1 - e^{-iq_1})}f_{1,1}^{(1)} \quad (E.4)
\]

We have again added the superscript (1) to distinguish \(f_{1,0}^{(1)}, f_{1,1}^{(1)}\) from the MA\((v, 0)\) expressions. With these definitions, the recurrence relations for the \(\delta f_{n,m}\) functions have precisely the same form as Eqs. (4.31)-(4.36) with \(n \to n - 1\) and \(\omega \to \omega - \Omega\). As before, we write this set of equations in the form \(\delta v_n = A_n \delta v_{n-1} + B_n \delta v_{n+1}\), where

\[
\delta v_n = \left(\delta f_{n,0}, \ldots, \delta f_{n,n-1}, \delta f_{n,0}, \ldots, \delta f_{n,n-2}, \delta f_{n,1}, \ldots, \delta f_{n,n-1}\right)^T, \quad (E.5)
\]
Appendix E. MA\textsuperscript{(v,1) Details}

and the solution can again be written in terms of an infinite continued fraction of matrices:

\[
\delta v_n = R_n \delta v_{n-1} = \frac{1}{1 - B_n' \frac{1}{1 - \ldots - A_n'}} \delta v_{n-1},
\]

(E.6)

where \( A_n' (\omega) = A_{n-1} (\omega - \Omega) \) and \( B_n' (\omega) = B_{n-1} (\omega - \Omega) \).

To finish the calculation we will require explicit expressions for \( f^{(1)}_{1,0}, f^{(1)}_{1,1}, f^{-1(1)}_{1,0}, \) and \( f^{+1(1)}_{1,1} \). Using the definitions in Eqs. (4.27), (4.28), (4.40), and (4.41) it is straightforward to show that

\[
f^{(1)}_{1,0} = \frac{1}{N} \sum_{q_1} G_0(k - q_1, \omega - \Omega) \frac{1}{1 - e^{iq_1}} \]

\[
\times \left[ \tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1) \right] + e_0 - e_1 e^{-ika},
\]

(E.7)

\[
f^{(1)}_{1,1} = -\frac{1}{N} \sum_{q_1} G_0(k - q_1, \omega - \Omega) \frac{1}{1 - e^{-iq_1}} \]

\[
\times \left[ \tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1) \right] - e_0 + e_1 e^{ika},
\]

(E.8)

\[
f^{-1(1)}_{1,0} = \frac{1}{N} \sum_{q_1} e^{i(k - q_1)} G_0(k - q_1, \omega - \Omega) \frac{1}{1 - e^{iq_1}} \]

\[
\times \left[ \tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1) \right] + e_1 - e_2 e^{-ika},
\]

(E.9)

and

\[
f^{+1(1)}_{1,1} = -\frac{1}{N} \sum_{q_1} e^{-i(k - q_1)} G_0(k - q_1, \omega - \Omega) \frac{1}{1 - e^{-iq_1}} \]

\[
\times \left[ \tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1) \right] - e_1 + e_2 e^{ika}.
\]

(E.10)

To combine all of these results and obtain a system of four equations in the four unknowns \( f^{(1)}_{1,0}, f^{(1)}_{1,1}, f^{-1(1)}_{1,0}, f^{+1(1)}_{1,1} \), we need to rewrite \( \tilde{f}_{2,0}(q_1) - \tilde{f}_{2,1}(q_1) \) in terms of these four quantities. Using the definitions in Eqs. (E.2)-(E.4),

165
Appendix E. MA($v,1$) Details

$v_2 = Q_2v_1$, and $\delta v_2 = R_1 \delta v_1$, we find that

\[
\bar{f}_{2,0}(q_1) - \bar{f}_{2,1}(q_1) = G_0^{-1}(k - q, \omega - \Omega)G_0(k - q, \tilde{\omega}) \times \left\{ (R_{00} - R_{10}) \left[ G_0(k - q_1, \omega - \Omega)g_{q_1}^2 - g(1 - e^{iq_1})f_{1,0}^{(1)} + g(1 - e^{-iq_1})f_{1,1}^{(1)} \right. \right.
\]

\[
\left. - ge^{-i(k-q_1)}(1 - e^{iq_1})f_{1,0}^{-(1)} + ge^{i(k-q_1)}(1 - e^{-iq_1})f_{1,1}^{+(1)} \right] + g(1 - e^{iq_1})Q_{2}v_{1|0} - 2g[1 - \cos(q_1)]Q_{2}v_{1|1} + g(1 - e^{-iq_1})Q_{2}v_{1|2} + ge^{-i(k-q_1)}Q_{2}v_{1|4} + ge^{i(k-q_1)}(1 - e^{-iq_1})Q_{2}v_{1|5} \right\}, \quad (E.11)
\]

where $\tilde{\omega} = \omega - \Omega - g(R_{00} - R_{10})$. In this notation $R_{ij}$ are the matrix elements of $R_1$ and $Q_2v_{1|i}$ is the $i^{th}$ element of the product between the matrix $Q_2$ and the vector $v_1$, i.e., $Q_{2}v_{1|i} = Q_{i0}f_{1,0}^{(1)} + Q_{i1}f_{1,1}^{(1)} + Q_{i2}f_{1,0}^{-(1)} + Q_{i3}f_{1,1}^{+(1)}$, where $Q_{ij}$ are the matrix elements of $Q_2$. Inserting Eq. (E.11) into Eqs. (E.7)-(E.10), and performing the required sums over momentum, we obtain the desired four equations in four unknowns. The explicit expressions for the momentum averages that appear in these expressions are given in Appendix D.

The resulting system of equations is easily solved, and the MA($v,1$) self-energy can be found from Eq. (4.29). The result is

\[
\Sigma_{MA^{(v,1)}}(k, \omega) = g(f_{1,0}^{(1)} - f_{1,1}^{(1)}). \quad (E.12)
\]
Appendix F

IMA\(^{(1)}\) Details

It is our goal to solve the recurrence relations of Eqs. (5.20) and (5.23) at the IMA\(^{(1)}\) level. We first illustrate how this is done for \(d = 1\).

We define the following Green’s functions:

\[
\bar{F}_n(i, j, j', \omega) = \langle 0 | c_i \hat{G}(\omega) c_j^{\dagger} b_j^{n-1} b_j'|0 \rangle. \quad (F.1)
\]

At the IMA\(^{(1)}\) level of approximation we set \(G_s(k - \mathbf{q}_T, j', j, \omega - n\Omega) \to 0\) for \(j \neq j'\) and \(n \geq 2\) in Eq. (5.23). Therefore, the \(n = 1\) case of Eq. (5.23) is kept exactly at the IMA\(^{(1)}\) level:

\[
F_1(i, j, \omega) = g \sum_{j \neq j'} G_s(j', j, \omega - \Omega) \bar{F}_2(i, j', j, \omega) + g G_s(j, j, \omega - n\Omega) [G(i, j, \omega) + F_2(i, j, \omega)]. \quad (F.2)
\]

and for \(n \geq 2\), we can use the IMA\(^{(0)}\) result [see Eq. (5.30)]:

\[
F_2(i, j, \omega) = g A_2(j, \omega) F_1(i, j, \omega). \quad (F.3)
\]

For \(j = j'\) we have

\[
\bar{F}_2(i, j, j, \omega) = F_2(i, j, \omega) = g A_2(j, \omega) F_1(i, j, \omega), \quad (F.4)
\]

and we also observe that direct analogy to the \(F_n\) recursion relations in Eq. (5.23) leads to

\[
\bar{F}_2(i, j, j', \omega) = g A_1(j, \omega - \Omega) \bar{F}_1(i, j, j', \omega). \quad (F.5)
\]

To proceed we generate an equation of motion for \(\bar{F}_1(i, j, j', \omega)\) using its definition in Eq. (F.1):

\[
\bar{F}_1(i, j, j', \omega) = g G_s(j', j, \omega - \Omega) G(i, j', \omega) + g \sum_{j''} G_s(j'', j, \omega - \Omega) \bar{F}_2(i, j'', j', \omega). \quad (F.6)
\]
Appendix F. IMA\(^{(1)}\) Details

After some manipulation, it can be shown that by substituting Eqs. (F.4) and (F.5) into the above equation of motion for \(\tilde{F}_1\), one can obtain the following:

\[
\tilde{F}_1(i, j, j', \omega) = gG_s(j', j, \tilde{\omega}_\Omega)\left[G(i, j', \omega) + \left[A_2(j', \omega) - A_1(j', \omega - \Omega)\right]\tilde{F}_1(i, j', j', \omega) + \sum_{j''} \tilde{F}_1(i, j'', j', \omega)v_0(j'', \omega - \Omega)G_s(j'', j, \tilde{\omega}_\Omega)\right],
\]

where \(\tilde{\omega}_\Omega = \omega - \Omega - g^2A_1(\omega - \Omega)\), and \(A_1(\omega)\) is defined in Eq. (5.33). Given the form of the inhomogeneous term in Eq. (F.7), we must have

\[
\tilde{F}_1(i, j, j', \omega) = gG_s(j', j, \tilde{\omega}_\Omega)\left[G(i, j', \omega) + \left[A_2(j', \omega) - A_1(j', \omega - \Omega)\right]\tilde{F}_1(i, j', j', \omega) + \sum_{j''} \tilde{F}_1(i, j'', j', \omega)v_0(j'', \omega - \Omega)G_s(j'', j, \tilde{\omega}_\Omega)\right],
\]

Inserting this “guess” into Eq. (F.7) and solving for \(x_{jj'}\), one finds that \(x_{jj'}\) has the same form as the IMA\(^{(0)}\) result of Eq. (5.35), but at a reduced frequency:

\[
x_{jj'}(\omega) = G_{IMA^{(0)}}(j, j', \omega - \Omega).
\]

Setting \(j = j'\) in Eq. (F.8), and using Eq. (F.4), we find that

\[
F_1(i, j, \omega) = \frac{gx_{jj}(\omega)}{1 - gx_{jj}(\omega)\left[A_2(j, \omega) - A_1(j, \omega - \Omega)\right]},
\]

which replaces the simpler IMA\(^{(0)}\) solution \(F_1(i, j, \omega) = A_1(j, \omega)G(i, j, \omega)\). This result can be substituted into Eq. (5.20), and we obtain the IMA\(^{(1)}\) solution

\[
G_{IMA^{(1)}}(i, j, \omega) = G_s(i, j, \omega) + \sum_{j' \geq 0} G_{IMA^{(1)}}(j, j', \omega)\Sigma_{IMA^{(1)}}(j', \omega)G_s(j', j, \omega),
\]

where \(\Sigma_{IMA^{(1)}}(j, \omega)\) is given by

\[
\Sigma_{IMA^{(1)}}(j, \omega) = \frac{g^2x_{jj}(\omega)}{1 - g^2x_{jj}(\omega)\left[A_2(j, \omega) - A_1(j, \omega - \Omega)\right]}.
\]
Since this solution has the same form as the IMA\(^{(0)}\) solution, we can again use the bulk result to speed up the convergence. Defining the effective interaction \(v_1(j, \omega) = \Sigma_{IMA^{(1)}}(j, \omega) - \Sigma_{MA^{(1)}}(\omega)\), and using the same trick as in the MA\(^{(0)}\) case, we arrive at the final IMA\(^{(1)}\) result for \(d = 1\) given in the main text [see Eq. (5.37)].

The result is easily be extended to higher dimensions. Carrying out the exact same steps as above, with the appropriate momentum labels and summations, one obtains the general solution given in Eq. (5.37), with \(x_{jj}\) replaced by its momentum average \(\bar{x}_{jj}\) [see Eq. (5.40)].
Appendix G

List of Publications

- Goodvin, G.L., Covaci, L., and Berciu, M., “The effect of a surface on a Holstein polaron”, has been submitted for publication.