ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DEPARTMENT OF PHYSICS AND ASTRONOMY "A. RIGHI"

SECOND CYCLE DEGREE

PHYSICS

TITLE OF DISSERTATION

Supervisor
Prof. Name Surname

Co-supervisor optional
Prof. Name Surname

Defended by
Name Surname

The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when one contemplates the mysteries of eternity, of life, of the marvellous structure of reality. It is enough if one tries to comprehend only a little of this mystery every day.

Albert Einstein

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Contents

List of Figures

List of Tables

1 Froehlich Polaron

1.1 Electrons in crystals

The study of electrons inside a solid crystal is an important field due to their role in the determination of transport and optical properties of such materials.

Multiple experiments involving X-ray or electron scattering from suitable solid samples have demonstrated that a crystalline solid, whether a metal, an insulator or a semiconductor, is described a periodic arrangement of atoms consisting of a unit cell (usually the primitive one, defined as the smallest unit cell possible in a determined material) with a suitable atomic basis: some examples are Iron (Fe), described by a BCC (Body-Centered Cubic) unit cell with a single Fe atom basis, rock salt (NaCl), described by by a FCC (Face-Centered Cubic) unit cell with a basis composed of one Na atom and one Cl atom, and crystalline silicon (Si), which is characterized by the diamond structure, an FCC structure with a basis composed of 2 silicon atoms (geometrically different from that seen in rock salt).

Given the periodic nature of crystal structures, it is natural to consider the electrostatic potential generated by the crystal to be periodic too. Such is the basic assumption which underlies the treatment of electrons in crystals, that was first described by Felix Bloch in his famous 1928 paper *Uber die Quantenmechanikder Elektronen in Kristallgittern* [6].

The model proposed by Bloch assumed independent electrons (no electron-electron interactions terms): each electron behaves as if only an average contribution to the potential from the other electrons (with the same periodicity of the lattice) exists, thus retrieving an **effective one-electron potential**.

The resulting Hamiltonian for the single independent electron is:

$$\mathrm{H}\psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \right)\psi(\mathbf{r}) = E\psi(\mathbf{r}),\tag{1.1}$$

where $V(r)$ is a periodic potential with the lattice periodicity, given a lattice translation vector

$$\mathbf{T} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \qquad n_1, n_2, n_3 \quad \text{integers},\tag{1.2}$$

with $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ primitive translation vectors. Therefore, we have

$$V(\mathbf{r} + \mathbf{T}) = V(\mathbf{r})\tag{1.3}$$

This means that the potential can be expanded in Fourier series

$$V(\mathbf{K}) = \sum_{\mathbf{K}} V_{\mathbf{K}} e^{i\mathbf{K}\cdot\mathbf{r}},\tag{1.4}$$

and is consequently useful to expand the wavefunction in the same way. For this reason the Born-von Karman boundary conditions are applied [7]

$$\psi(\mathbf{r} + N_i\mathbf{a}_i) = \psi(\mathbf{r}), \qquad i = 1, 2, 3.\tag{1.5}$$

where $N_i \mathbf{a}_i$ is a chosen vector such that $V(r + N_i a_i) = V(\mathbf{r})$ (it is a translation vector for the lattice). This is a normalizing condition for the wavefunction which corresponds to taking N_i unit cell in each direction and periodically repeating them. In this way it is possible to write the wavefunction in the following way:

$$\psi(\mathbf{r}) = \sum_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}. \tag{1.6}$$

Going back to the periodic potential $U(\mathbf{r})$ we observe that the Fourier terms $U_{\mathbf{K}}$ are related to $U(\mathbf{r})$ by

$$U_{\mathbf{K}} = \frac{1}{V} \int_{cell} dr e^{-i\mathbf{K}\cdot\mathbf{r}} U(\mathbf{r}) \tag{1.7}$$

with V unit cell volume and $\mathbf{K}$ reciprocal lattice vector (vector in k-space such that $e^{i\mathbf{K}\cdot\mathbf{a}_i} = 1$). If it is assumed that the potential $V(\mathbf{r})$ is real and $V(-\mathbf{r}) = V(\mathbf{r})$ (valid for every Bravais lattice) it follows that the coefficients $U_{\mathbf{K}}$ are real.
Substituting in 1.1 we obtain:

$$(\mathrm{H} - E)\psi(\mathbf{r}) = \sum_{\mathbf{q}} \left(\frac{\hbar^2}{2m} q^2 - E \right) c_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} + \sum_{\mathbf{K}\mathbf{q}'} U_{\mathbf{K}} c_{\mathbf{q}'-\mathbf{K}} e^{i\mathbf{q}'\cdot\mathbf{r}} = 0, \tag{1.8}$$

which yields the following result:

$$\left(\frac{\hbar^2}{2m} (\mathbf{k} - \mathbf{K})^2 - E \right) c_{\mathbf{k}-\mathbf{K}} + \sum_{\mathbf{K}'} c_{\mathbf{K}'-\mathbf{K}} c_{\mathbf{k}-\mathbf{K}'} = 0. \tag{1.9}$$

From this equation it becomes clear that for a fixed k only the coefficients $c_{\mathbf{k}}, c_{\mathbf{k}-\mathbf{K}}, \dots$ whose wavevector differs from $\mathbf{k}$ by a reciprocal lattice vector are coupled. In this way the original problem has been divided in N independent equations for each allowed value of $\mathbf{k}$ (in the first Brillouin zone, the primitive unit cell in the reciprocal space).
Going back to the expansion 1.6 the wavefunction now becomes:

$$\psi_k(r) = \sum_{K} c_{k-K} e^{i(k-K)\cdot r}, \tag{1.10}$$

which can be recast as

$$\psi_k(r) = e^{ik\cdot r} \sum_{K} c_{k-K} e^{-iK\cdot R}; \tag{1.11}$$

in this form it is easy to recognize that

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{K}} c_{\mathbf{k}-\mathbf{K}} e^{-i\mathbf{K}\cdot\mathbf{R}} \tag{1.12}$$

has the periodicity of the (reciprocal) Bravais lattice.

The shape of the Bloch wavefunction as stated in the Bloch theorem is thus retrieved

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) \tag{1.13}$$

and its square modulus (from which important properties such as charge distribution follow) $|\psi_{\mathbf{k}}(\mathbf{r})|^2$ has the periodicity of the lattice.

The new Hamiltonian eigenvalues equation using Bloch wavefunction reads

$$\mathrm{H}_{\mathbf{k}}u_{n\mathbf{k}} = \left(\frac{\hbar^2}{2m} \left(\frac{1}{i}\nabla + \mathbf{k} \right)^2 + U(\mathbf{r}) \right) u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}), \tag{1.14}$$

which yields a different energy eigenvalue for each $\mathbf{k}$ in the first Brillouin zone. The index n arises from the fact that for each $\mathbf{k}$ value infinite discrete solutions to the eigenvalue equation exist, in the same as in the particle in a box problem. Going to the thermodynamic limit ($N_i \to \infty$ for each component i) the entire first Brillouin zone is sampled and the typical band structure of the solids is recovered.

All the methods employed in solid state calculations rely on this "simple" equation to determine band structures, granted analytical form of the periodic potential $U(\mathbf{r})$ can be retrieved. The most used and effective computation method is **Density Functional Theory** [8] or **DFT** for short, which does not directly compute the eigenvalue equation for each single k and n value, which would be impossible since the potential $U(\mathbf{r})$ is a many-body object that includes electron-electon interactions, but rather computes all the physical quantities of interest starting from the ground state charge density $\rho_0(\mathbf{r})$. This is made possible by the two Hohenberg-Kohn theorems [9], which respectively state that the external potential (the periodic $U(\mathbf{r})$ from the electron point of view) is uniquely determined up to a constant by the ground state charge density $\rho_0(\mathbf{r})$ and that this density can be computed variationally by minimizing a suitable functional $F[\rho]$.

Nevertheless, the functional $F[\rho]$ exact form is unknown: for this reason the **Kohn-Sham scheme** [10] is used. The idea at the basis of this practical setup is to substitute the real interacting system (many-body) with an auxiliary non-interacting one (single particle) which has the same ground state as the original. In this system the total potential $V(\mathbf{r}) = V_{ion}(\mathbf{r}) + V_{ee}(\mathbf{r})$ is replaced with an effective potential $V_{eff}(\mathbf{r})$ defined as

$$U_{eff}(\mathbf{r}) = U_{ion}(\mathbf{r}) + U_H(\mathbf{r}) + U_{xc}(\mathbf{r}), \tag{1.15}$$

where U_{ion} is the ionic potential, often modelled using pseudopotentials [11], $U_H(\mathbf{r})$ is the Hartree term, the electrostatic potential due to the mean field distribution of all the other electrons, and $U_{xc}(\mathbf{r})$ the exchange-correlation term, an approximate potential which includes Pauli exclusion principle (exchange), higher order electron-electron interactions (correlation) and many body corrections to kinetic energy. The exchange-correlation term is usually computed using **LDA** (Local Density Approximation) or **GGA** (Generalized Gradient Approximation),

both based on the the Uniform Electron Gas (UEG) [12]. Despite the inevitable approximations, DFT calculations are able to provide extremely accurate band structure calculations with few drawbacks, such as inaccurate gap estimations in insulators and semiconductors.
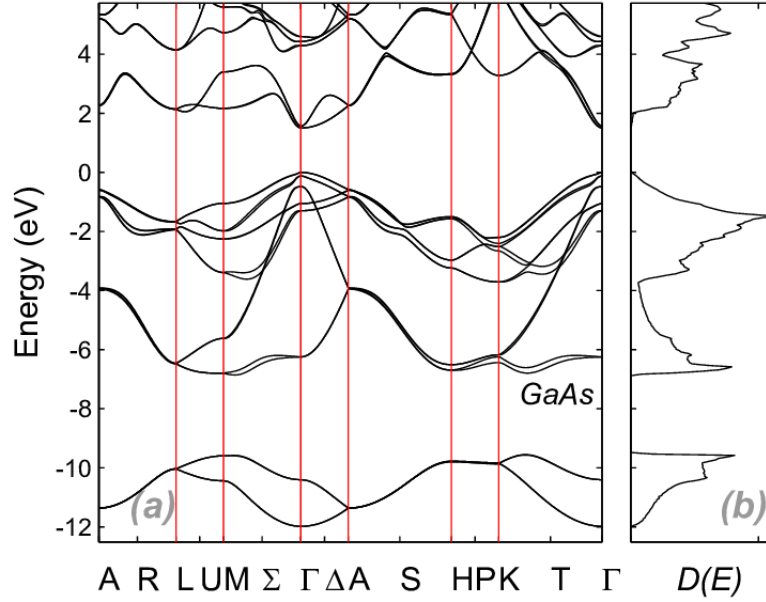
Figure 1.1: DFT calculated low-temperature GaAs Wurtzite structure with relative DOS [1].

As it can be seen in Figure 1.1 for GaAs, at the band extrema (at the edge between the allowed band states and the forbidden gap) the bands have a quadratic dispersion akin to that of a free electron: an electron in these regions only feel a very weak external potential $U(\mathbf{r})$ and can be considered as free provided that the free electron mass m is substituted with an effective mass m^* which takes into account the potential acting on it.
In a general case the effective mass is a tensorial quantity (bands are anisotropic in general with respect to the wavevector) and their value can be computed with

$$m_{ij}^{*-1} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j},\tag{1.16}$$

provided that the quadratic approximation is valid. This approximation is really useful when calculating quantities near band extrema, providing an effective description of an otherwise complex interaction.

1.2   The polaron problem: the Froehlich model

The treatment described up until now for electrons in solids relies on a basic assumption: electro dynamics is much faster than ion dynamics. This is the so-called **Born-Oppenheimer**

approximation, which assumes that the motion of the much heavier ions is slow compared to the fast dynamics of the electrons (the characteristic time are of the order of ps for ions and fs for electrons). As a result, electrons are considered to respond instantaneously to changes in the positions of the ions, allowing the separation of electronic and ionic degrees of freedom in the analysis. Given M ions and n electrons the wavefunction $\Psi(\mathbf{R}_1, ..., \mathbf{R}_M, \mathbf{r}_1, ..., \mathbf{r}_n)$ with $\mathbf{R}_\alpha$ ionic positions and $\mathbf{r}_i$ electron positions can be decomposed in the following way:

$$\Psi(\mathbf{R}_1, ..., \mathbf{R}_M, \mathbf{r}_1, ..., \mathbf{r}_n) = \chi(\mathbf{R}_1, ..., \mathbf{R}_M)\psi(\mathbf{r}_1, ..., \mathbf{r}_n), \tag{1.17}$$

and only the electronic wavefunction $\psi(\mathbf{r}_1, ..., \mathbf{r}_n)$ is taken into consideration, decoupling ionic degrees of freedom and treating ionic potential as a constant energy surface $U_{ion}(\mathbf{r})$.

This approximation, although extremely effective in computing electronic bands, fails by design to describe systems where there is a strong coupling between electrons and ions (for example electron-phonon interaction).

Such is the case of polar crystals: in these materials single electrons at the bottom of the conduction band (or holes at the top of the valence band) couple to the strongly polarized ions and distort the unit cell: an electron-phonon interaction is thus present in the lattice and the electron properties (such as energy and effective mass) are renormalized due to this coupling [13].

The basic assumption of the Froehlich model is that the polaron is "large", namely the characteristic size of the polaron (electron together with its phonon cloud) is much larger than the lattice constant a. In this way it is possible to ignore the atomic details of the actual material and treat it as a uniform dielectric medium.

It is assumed that the electron which takes part in the formation of the polaron lies at the bottom of the valence band, in this way it is expected to have a quadratic energy dispersion characterized (as already seen in 1.16) by an effective mass

$$\epsilon(k) = \frac{\hbar^2 k^2}{2m^*}, \tag{1.18}$$

where the effective mass m^* encapsulates all the interactions due to crystal structure (for a hole polaron the dispersion is similar considering a negative effective mass).

Given the fact that we are dealing with the large polaron model, it is possible to treat the material as an homogeneous continuum which is polarized by the excess charge (the conduction electron). We also assume that the material is isotropic. Let $\mathbf{P}(\mathbf{r})$ be the electric polarization at the point $\mathbf{r}$, the total electric field is then:

$$\begin{aligned}
\mathbf{D}(\mathbf{r}) &= \epsilon_0 \mathbf{E}(\mathbf{r}) + 4\pi \mathbf{P}(\mathbf{r}), \\
\mathbf{D}(\mathbf{r}) &= \epsilon \mathbf{E}(\mathbf{r}).
\end{aligned} \tag{1.19}$$

The only free charge present in the material is the conduction electron in excess, this means that $\mathbf{D}(\mathbf{r})$ is completely determined by it:

$$\begin{aligned}
D(\mathbf{r}, \mathbf{r}_{el}) &= -\nabla \frac{e}{|\mathbf{r} - \mathbf{r}_{el}|}, \\
\nabla \cdot \mathbf{D}(\mathbf{r}, \mathbf{r}_{el}) &= 4\pi\delta(\mathbf{r} - \mathbf{r}_{el}),
\end{aligned} \tag{1.20}$$

with $\mathbf{r}_{el}$ the coordinate of the conduction electron. Developing the equations the following relation is obtained:

$$4\pi\mathbf{P}(\mathbf{r}) = \left(1 - \frac{1}{\epsilon}\right)\mathbf{D}(\mathbf{r}).\tag{1.21}$$

$\mathbf{P}(\mathbf{r})$ is the total polarization, which has the contribution both from electrons and from ionic displacements, since only the second contribution is relevant in our treatment, a method to tell apart this two contributions is needed.

To obtain this we imagine a system where a field is slowly applied and then rapidly switched off [14]: only electrons are able to keep up with the variation of the field. A formula connecting polarization and $\mathbf{D}$ (dropping the $\mathbf{r}$ dependence for simplicity since an isotropic and homogeneous medium is assumed) is then retrieved as

$$4\pi\delta\mathbf{P} = -\left(1 - \frac{1}{\epsilon_\infty}\right)\mathbf{D},\tag{1.22}$$

it is then obtained the lattice contribution to the polarization:

$$4\pi\mathbf{P}_{lat} = 4\pi(\mathbf{P} - \delta\mathbf{P}) = \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\mathbf{D}.\tag{1.23}$$

We know consider a model for the polar crystal lattice: we assume it to be constituted of individual oscillating discrete dipoles located at fixed sites of the crystal with characteristic frequency ω. The energy of each one of these oscillators is

$$\frac{M}{2}(\dot{q}^2 + \omega^2 q^2).\tag{1.24}$$

The next step is to move from a discrete set of fixed oscillators to a continuum of dipoles: it is assumed that the dipoles are not coupled. Defining the effective charge as e^*, the following substitution is made:

$$e^*\mathbf{q}_n \to \mathbf{P}_{lat}(\mathbf{r}),\tag{1.25}$$

which relates the dipole moment to the polarization field. We also write the following relation:

$$\frac{M}{e^{*2}} = \gamma d^3 r,\tag{1.26}$$

where $M = nd^3 r$ is the mass density of the medium.

It is now possible to define the kinetic and potential energy of the freely oscillating polarization field:

$$T = \int \frac{\gamma}{2}\dot{\mathbf{P}}_{lat}^2(\mathbf{r})d^3 r,$$
$$V = \int \frac{\gamma\omega^2}{2}\mathbf{P}_{lat}^2(\mathbf{r})d^3 r.\tag{1.27}$$

We identify ω with the optical phonon frequency ω_{LO}, but the constant γ has not yet been given a physical definition. For this reason we consider the interaction between electrons and dipole moment, which is given by

$$e\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\mathbf{P}_{lat},$$ (1.28)

which, using a continuous charge distribution and the polarization field, becomes

$$\rho(\mathbf{r})d^3r\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\mathbf{P}_{lat}(\mathbf{r}')d^3r'.$$ (1.29)

The interaction energy E_I is thus

$$E_I = \iint \rho(\mathbf{r})\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\mathbf{P}_{lat}(\mathbf{r}')d^3r d^3r'.$$ (1.30)

The Lagrangian is then built as follows:

$$L = T - U - E_I$$ (1.31)

and the equation of motion are found solving the Lagrange equations

$$\frac{d}{dt}\frac{\delta L}{\delta \dot{P}_{lat_i}} - \frac{\delta L}{\delta P_{lat_i}} = 0$$ (1.32)

for $i = 1, 2, 3$. The explicit formula of the equation of motion is:

$$\gamma(\ddot{\mathbf{P}}_{lat}(\mathbf{r}') + \omega^2\mathbf{P}_{lat}(\mathbf{r}')) = -\int \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}\rho(\mathbf{r})d^3r$$ (1.33)

keeping fixed $\rho(\mathbf{r})$. The term on the right side is simply the dielectric displacement $\mathbf{D}(\mathbf{r}')$ and, in the static limit, the equation simplifies to

$$\gamma\omega^2\mathbf{P}_{lat}(\mathbf{r}') = \mathbf{D}(\mathbf{r}').$$ (1.34)

Comparing this equation with 1.23, the value of γ is obtained:

$$\gamma = \frac{4\pi}{\omega^2}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)^{-1}.$$ (1.35)

It is now time to define the full complete polaron Hamiltonian. We start by defining the free electron term (the interaction lies inside the effective mass m^*), considering that the electron charge density operator is defined as:

$$\rho(\mathbf{r}) = -e\psi^\dagger(\mathbf{r})\psi(\mathbf{r})$$ (1.36)

using the second quantization formalism with field operators, then the electron kinetic term is

$$H_{el} = \int \psi^\dagger(\mathbf{r}) \left(-\frac{\hbar^2}{2m^*} \nabla^2 \right) \psi(\mathbf{r}) d^3r. \tag{1.37}$$

If we use a suitable basis for the field operator

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} c_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1.38}$$

with V (unit cell) volume, then 1.37 becomes

$$H_{el} = \sum_{\mathbf{k}} \frac{k^2}{2m^*} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}, \tag{1.39}$$

with $c_{\mathbf{k}}^\dagger$ and $c_{\mathbf{k}}$ creation and annihilation operators for the free electron.
We now turn our attention to the polarization field, taking into account that it can be modelled as an harmonic oscillator a form similar to 1.37 can be recovered (P_{lat} is substituted by P in the following equations):

$$H_{ph} = \int \left(\frac{1}{2\gamma} \mathbf{\Pi}^2(\mathbf{r}) + \frac{\gamma}{2}\omega^2 \mathbf{P}^2(\mathbf{r}) \right) d^3r, \tag{1.40}$$

where $\mathbf{\Pi}(\mathbf{r})$ is the conjugate momentum to $\mathbf{P}(\mathbf{r})$.
Using the canonical substitution for the operator $\mathbf{P}(\mathbf{r})$

$$\mathbf{P}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} \sqrt{\frac{\hbar}{2\gamma\omega}} \frac{\mathbf{q}}{q} e^{i\mathbf{q}\cdot\mathbf{r}} (a_{\mathbf{q}} - a_{-\mathbf{q}}^\dagger) \tag{1.41}$$

and the corresponding one for $\mathbf{\Pi}(\mathbf{r})$ we obtain the equation for a quantum harmonic oscillator:

$$H_{ph} = \hbar\omega \sum_{\mathbf{q}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}}, \tag{1.42}$$

with $a_{\mathbf{q}}^\dagger$ and $a_{\mathbf{q}}$ bosonic phononic creation and annihilation operators.
Using the same formalism, the interaction term is instead written in the following way:

$$H_{EPC} = \iint \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \frac{e}{|\mathbf{r} - \mathbf{r}'|} (-\nabla_{\mathbf{r}'} \cdot \mathbf{P})(\mathbf{r}') d^3r d^3r', \tag{1.43}$$

using the same substitution in 1.41 for $\mathbf{P}(\mathbf{r})$ the following form is obtained:

$$H_{EPC} = \int \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \left[4\pi i \left(\frac{e^2\hbar}{2\gamma\omega V} \right)^{1/2} \sum_{\mathbf{q}} \frac{1}{q} (a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}} - a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}) d^3r \right]. \tag{1.44}$$

The term is simplified even more if we substitute the electron field operator using 1.41, then the final form of the coupling term is reached

$$H_{EPC} = \sum_{\mathbf{k},\mathbf{q}} (V_{\mathbf{w}} c^\dagger_{\mathbf{k+q}} c_{\mathbf{k}} a_{\mathbf{w}} + V^*_{\mathbf{w}} c^\dagger_{\mathbf{k-q}} c_{\mathbf{k}} a^\dagger_{\mathbf{q}}), \tag{1.45}$$

where the coupling parameter is given as

$$V_{\mathbf{q}} = -4\pi i \left(\frac{e^2}{2\hbar \gamma \omega V} \right)^{1/2} \frac{1}{q}. \tag{1.46}$$

It is possible to rewrite the coupling parameter in terms of the dimensionless coupling strength α:

$$V_{\mathbf{q}} = i \left(\frac{4\pi\alpha}{V} \right)^{1/2} \frac{1}{q},$$
$$\alpha = \frac{2\pi e^2}{\hbar \gamma \omega} \sqrt{\frac{2m^*\omega}{\hbar}} = \frac{1}{2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) \frac{e^2}{\hbar \omega} \sqrt{\frac{2m^*\omega}{\hbar}}. \tag{1.47}$$

If we now identify ω with ω_{LO} angular frequency of the longitudinal optical phonon modes (the modes that effectively couple with the electron), we obtain the final form of the Froehlich Hamiltonian:

$$H^{Fr} = \sum_{\mathbf{k}} \frac{k^2}{2m^*} c^\dagger_{\mathbf{k}} c_{\mathbf{k}} + \hbar \omega_{LO} \sum_{\mathbf{q}} a^\dagger_{\mathbf{q}} a_{\mathbf{q}} + \sum_{\mathbf{k},\mathbf{q}} (V_q c^\dagger_{\mathbf{k+q}} c_{\mathbf{k}} a_{\mathbf{q}} + V^*_q c_{\mathbf{k-q}} c_{\mathbf{k}} a^\dagger_{\mathbf{q}}) \tag{1.48}$$

where we have dropped the coupling parameter dependence on $\mathbf{q}$ direction since it only depends on its modulus.
The main assumptions of this model are:

- Free electron with quadratic dispersion (up to a *scalar* effective mass m^*).

- Dispersionless optical mode with frequency ω_{LO}.

- "Large" polaron, with characteristic dimension d much greater than lattice parameter a (continuum approximation).

In this form the model correctly describes only a handful of real materials and it mainly serves as a toy model (we will breafly mention how it can be solved at weak coupling with perturbation theory and at strong coupling variationally), however we will focus on the **Diagrammatic Monte Carlo Method**, a computational method that is capable of solving the polaron problem using some tricks. We will also see how some of the main limitations of the model (namely the scalar effective mass m^* and the single phonon mode ω_{LO}) can be lifted without lifting the main assumption (the large polaron approximation). It will be illustrated in particular the case of an anisotropic band with effective mass $m^*(\hat{\mathbf{k}})$ and multiple dispersionless optical modes ω_{LO_j}, already capable of modelling valence band minima in a wide range of materials.

1.3   Weak coupling limit: perturbation theory

In the weak coupling regime it is possible to apply perturbation theory to the Froehlich Hamiltonian, for this treatment a slightly modified Hamiltonian with respect to the one seen in 1.48 will be used [15]:

$$H^{Fr} = -\frac{\nabla^2}{2m^*} + \omega_{LO} \sum_{\mathbf{q}} (a_{\mathbf{q}}^\dagger a_{\mathbf{q}}) + \sum_{\mathbf{q}} (V_{\mathbf{q}} a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} + V_{\mathbf{q}}^* a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}}). \tag{1.49}$$

We will respectively define the non-interacting Hamiltonian H_0^{Fr} and the interaction term H_I^{Fr} as

$$
\begin{aligned}
H_0^{Fr} &= -\frac{\nabla^2}{2m^*} + \omega_{LO} \sum_{\mathbf{q}} (a_{\mathbf{q}}^\dagger a_{\mathbf{q}}), \\
H_I^{Fr} &= \sum_{\mathbf{q}} (V_{\mathbf{q}} a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} + V_{\mathbf{q}}^* a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}}).
\end{aligned}
\tag{1.50}
$$

The quantum state of the non-interacting electron is described by $|\mathbf{k}\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}$, while for the non interacting phonons it is possible to define the average number of excitations $\langle a_{\mathbf{q}}^\dagger a_{\mathbf{q}} \rangle = \langle n_{\mathbf{q}} \rangle$, which is equal to 0 at the ground state The total wavefunction of the non-interacting system then becomes:

$$|\mathbf{k}, 0\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} |0\rangle. \tag{1.51}$$

The first excited state to which the ground non-interacting state can jump to is the one with electron wavevector $\mathbf{k} - \mathbf{q}$ and one phonon $n_{\mathbf{q}} = 1$, total energy is:

$$\frac{(\mathbf{k} + \mathbf{q})^2}{2m^*} + \omega_{LO} = \frac{k^2}{2m^*}, \tag{1.52}$$

the corresponding wavefunction is

$$|\mathbf{k} - \mathbf{q}, 1\rangle = e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}} |1\rangle. \tag{1.53}$$

At first order perturbation theory there is no correction to the energy, we then need to go to second order. We consider the matrix element

$$\langle \mathbf{k} - \mathbf{q}, 1 | H_I^{Fr} | \mathbf{k}, 0 \rangle = V_{\mathbf{q}}^*, \tag{1.54}$$

second order perturbation theory energy then becomes:

$$E_{\mathbf{k}P} = \frac{k^2}{2m^*} - \sum_{\mathbf{q}} \frac{|V_{\mathbf{q}}|^2}{(\mathbf{k} - \mathbf{q})^2/2m^* + \omega_{LO} - k^2/2m^*}, \tag{1.55}$$

evaluating the sum over the q wavevectors we arrive at

$$E_{\mathbf{k}P} = \frac{k^2}{2m^*} - \frac{\alpha\omega_{LO}\sqrt{2m^*\omega_{LO}}}{k}\arcsin\frac{k}{\sqrt{2m^*\omega_{LO}}},\tag{1.56}$$

which for $k \ll \sqrt{2m^*\omega_{LO}}$ yields

$$E_{\mathbf{k}P} = \frac{k^2}{2m^*} - \alpha\omega_{LO}.\tag{1.57}$$

The polaron effective mass m_P^* is given by:

$$m_P^* = \frac{m^*}{\left(1 - \frac{\alpha}{6}\right)}.\tag{1.58}$$

From the equation it clearly results that the polaron effective mass diverges for $\alpha \to 6^-$ and is negative for $\alpha > 6$. This clearly signals that perturbation theory does not correctly describe the physical system at strong couplings.

1.4   Strong coupling limit: variational treatment

For the strong coupling case a different approach is taken: the main assumption of this model is a localized polaron wavefunction with a gaussian form [16], the Froehlich Hamiltonian is recast in order to write the phonon operators as displacement and conjugate moment operators:

$$Q_{\mathbf{q}} = \frac{1}{\sqrt{2}}\left(a_{\mathbf{q}} + a_{-\mathbf{q}}^\dagger\right), \qquad P_{\mathbf{q}} = -\frac{i}{\sqrt{2}}\left(a_{\mathbf{q}} - a_{-\mathbf{q}}^\dagger\right),\tag{1.59}$$

so that the Hamiltonian in 1.48 can be rewritten as

$$H^{Fr} = \frac{p^2}{2m^*} + \frac{\omega_{LO}}{2}\sum_{\mathbf{q}}\left(P_{\mathbf{q}}^2 + Q_{\mathbf{q}}^2\right) + \sum_{\mathbf{q}}V_{\mathbf{q}}Q_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{r}}.\tag{1.60}$$

A wavefunction with dependence both on the electron position $\mathbf{r}$ and the lattice displacement $Q_{\mathbf{q}}$ is needed, for this reason the following gaussian ansatz is used:

$$\Phi(\mathbf{r}, Q_{\mathbf{q}}) = \phi(\mathbf{r})\Psi_n(Q_{\mathbf{q}} + \delta Q_{\mathbf{q}}),$$
$$\phi(r) = \left(\frac{\beta}{\sqrt{\pi}}\right)^{3/2}e^{-\frac{\beta^2}{2}r^2},\tag{1.61}$$

with β variational parameter, Ψ_n the harmonic oscillator wavefunction and $\delta Q_{\mathbf{q}}$ displacement to be calculated.
We now need to take the expectation value of the Hamiltonian operator:

$$H(Q_{\mathbf{q}}) = \langle\phi(r)|H|\phi^*(r)\rangle = \int \phi(r)H\phi(r)d^3r,\tag{1.62}$$

the terms with $\mathbf{r}$ dependence are respectively evaluated as:

$$
\int \phi^*(\mathbf{r}) \frac{p^2}{2m^*} \phi(\mathbf{r}) d^3 r = \frac{3\beta^2}{4m^*},
$$
$$
\int \phi^*(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \phi(\mathbf{r}) d^3 r = e^{-q^2/4\beta^2}.
$$
(1.63)

The following result is thus obtained:

$$
H(Q_\mathbf{q}) = \frac{3\beta^2}{4m^*} + \frac{\omega_{LO}}{2} \sum_\mathbf{q} (P_\mathbf{q}^2 + Q_\mathbf{q}^2) + \sum_\mathbf{q} L_\mathbf{q} Q_\mathbf{q}, \qquad L_\mathbf{q} = V_\mathbf{q} e^{-q^2/4\beta^2}.
$$
(1.64)

To cancel out the linear term in $Q_\mathbf{q}$ it is important to choose the right equilibrium displacement $\delta Q_\mathbf{q}$ as

$$
\delta Q_\mathbf{q} = \frac{L_\mathbf{q}}{\omega_{LO}},
$$
(1.65)

which yields the Hamiltonian

$$
\begin{aligned}
H(Q_\mathbf{q}) &= \frac{\omega_{LO}}{2} \sum_\mathbf{q} \left[P_\mathbf{q}^2 + (Q_\mathbf{q} + \delta Q_\mathbf{q})^2 \right] + \frac{3\beta}{4m^*} - \frac{1}{2\omega_{LO}} \sum_\mathbf{q} L_\mathbf{q}^2 \\
&= \frac{\omega_{LO}}{2} \sum_\mathbf{q} \left[P_\mathbf{q}^2 + (Q_\mathbf{q} + \delta Q_\mathbf{q})^2 \right] + E(\beta).
\end{aligned}
$$
(1.66)

The three different terms have precise physical meanings: the first term describes the harmonic oscillation of phonons around their new equilibrium position $(Q_\mathbf{q} + \delta Q_\mathbf{q})$, the second term represents the kinetic energy of the electron in the gaussian formalism, the third one the interaction energy between the electron and the phonons. To variationally retrieve the lowest energy the β parameter is minimized, we have:

$$
\frac{1}{2\omega_{LO}} \sum_\mathbf{q} L_\mathbf{q}^2 = \alpha \left(\frac{\beta^2 \omega_{LO}}{m^* \pi} \right)^{1/2},
$$
$$
E(\beta) = \frac{3\beta^2}{4m^*} - \alpha \left(\frac{\beta^2 \omega_{LO}}{m^* \pi} \right)^{1/2}.
$$
(1.67)

If we calculate the derivative with respect to β and we minimize it the following result is obtained:

$$
\frac{dE}{d\beta} = \frac{3\beta}{2m^*} - \alpha \left(\frac{\omega_{LO}}{m^* \pi} \right)^{1/2} = 0,
$$
(1.68)

which yields a value for β_0:

$$
\beta_0 = \frac{\alpha}{3} 2m^* \left(\frac{\omega_{LO}}{m^* \pi} \right)^{1/2}.
$$
(1.69)

Using this result the minimum energy $E(\beta_0)$ is:

$$E(\beta_0) = -\frac{\alpha^2 \omega_{LO}}{3\pi} = -0.106\alpha^2 \omega_{LO}, \qquad (1.70)$$

with a quadratic dependence on the coupling strength α different from that found using perturbation theory (linear).

A more refined treatment of the strong coupling limit [17] yields the following result:

$$\lim_{a \to \infty} E_0(\alpha) = -\omega_{LO}\left[-0.1085\alpha^2 + 2.836 + O(1/\alpha^2)\right], \qquad (1.71)$$

not so different from our obtained result.
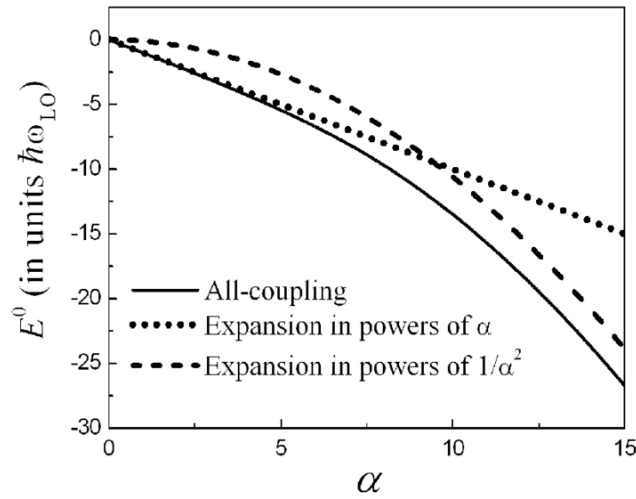
Figure 1.2: Computed polaron energy using perturbation theory, strong coupling theory and all coupling Feynman technique [2].

1.5   Froehlich Hamiltonian: degenerate bands and multiple phonon modes

The Hamiltonian in 1.48 only describes an extremely limited number of real cases, namely a non-degenerate isotropic electron band with effective mass m^* coupled to a single LO phonon mode. Although this relatively simple model is already quite challenging, some of its assumptions can be relaxed to obtain a more general framework.

Here we will be describing the **generalized Froehlich cubic model**: while mantaining some core simplification already seen in the original Froehlich model such as an isotropic dielectric tensor ϵ_0 and ϵ_∞ together with a phonon dispersion ω_{jLO} which does not depend on the wavevector direction (an even more general model not restrained to the cubic case is described in [18]) $\hat{k}$, it is still a quite relevant and useful model since it can be used to describe

multiple materials such as various oxides (BaO, CaO, MgO), II-VI compounds (CdS, CdSe, ZnS, ZnTe) and III-V compounds (AlAs, GaAs, GaN, GaP) [19].
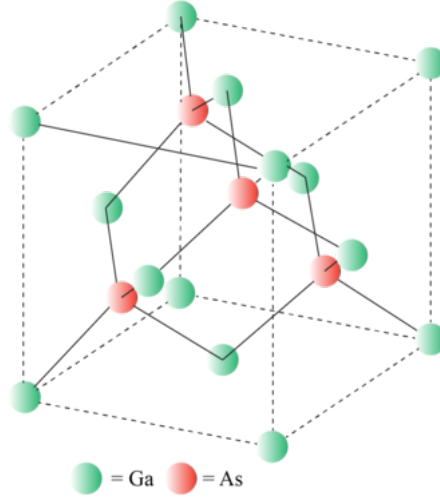
Figure 1.3: GaAs crystal structure, the cubic structure of the zincblende unit cell is here clearly displayed.

We start by defining the new electron term:

$$H_{el}^{cFr} \sum_{\mathbf{k}n} \frac{\sigma k^2}{2m_n^*(\hat{k})} c_{\mathbf{k}n}^\dagger c_{\mathbf{k}n}, \tag{1.72}$$

where the sum has been extended to n bands degenerate at the extremum, $\sigma = \pm 1$ depending on the fact that an electron or a hole polaron is taken into consideration, $m_n^*(\hat{k})$ is the new effective mass depending on band index n and wavevector direction $\hat{k}$ (a tensorial quantity in general).

We also want to extend our Hamiltonian to tackle the case of multiple LO phonon modes present in the material, to this aim we define the following phonon term:

$$H_{ph}^{cFr} = \sum_{\mathbf{q}j} \hbar \omega_{jLO} a_{\mathbf{q}j}^\dagger a_{\mathbf{q}j}, \tag{1.73}$$

where the index j represents the different LO modes.
The electron-phonon coupling term gets redefined as:

$$H_{EPC}^{cFr} = \sum_{\mathbf{k}nn',\mathbf{q}j} c_{\mathbf{k}+\mathbf{q}n'}^\dagger c_{\mathbf{k}n} \left[V^{cFr}(\mathbf{k}nn', \mathbf{q}j) a_{\mathbf{q}j} + V^{*cFr}(\mathbf{k}nn', \mathbf{q}j) a_{-\mathbf{q}j}^\dagger \right]. \tag{1.74}$$

It should be noted that the new electron-phonon coupling term includes interband transitions (from n to n'). The new coupling strength $V^{cFr}(\mathbf{k}nn', \mathbf{q}j)$ is found to be

$$V^{cFr}(\mathbf{k}nn', \mathbf{q}j) = \frac{i}{q}\frac{4\pi}{\Omega_0}\left(\frac{1}{2\omega_{jLO}V_{BvK}}\right)^{1/2}\frac{p_{jLO}}{\epsilon_\infty}\times\sum_m s_{n'm}(\hat{k}')s^*_{nm}(\hat{k}), \qquad (1.75)$$

with Ω_0 volume of the primitive unit cell, V_{BvK} volume of the Born-von Karman unit cell, p_{jLO} phonon mode polarities, linked to Born effective charges and ionic dielectric tensor ϵ_0 [20] and computed as $p_{jLO}(\hat{q}) = \sum_\nu Z^*_\nu e_\nu(\hat{q})$ isotropic in cubic system, and the s tensors, symmetry dependent unitary matrices which represent specific bands (and have the same symmetry group).
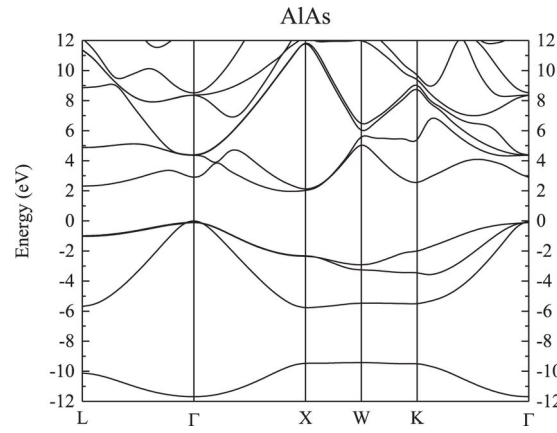
Figure 1.4: DFT calculated AlAs band structure [3], the conduction band minimum is on the high symmetry line Δ between Γ and X.

In our case of interest we restrict to just one anisotropic non-degenerate electron band, which is an accurate model for the conduction band minimum of the materials cited before (usually located somewhere between Γ and X along the Δ line), in this case the interaction strength simplifies to:

$$V^{cFr}(qj) = \frac{i}{q}\frac{4\pi}{\Omega_0}\left(\frac{1}{2\omega_{jLO}V_{BvK}}\right)^{1/2}\frac{p_{jLO}}{\epsilon_\infty}, \qquad (1.76)$$

and the full Hamiltonian reads:

$$H^{cFr} = \sum_\mathbf{k}\frac{\sigma k^2}{2m^*(\hat{k})}c^\dagger_\mathbf{k}c_\mathbf{k} + \sum_{\mathbf{q}j}\hbar\omega_{jLO}a^\dagger_{\mathbf{q}j}a_{\mathbf{q}j} + \sum_{\mathbf{k},\mathbf{q}j}c^\dagger_{\mathbf{k}+\mathbf{q}}c_\mathbf{k}\left[V^{cFr}(qj)a_\mathbf{q} + V^{*cFr}(qj)a^\dagger_{-\mathbf{q}}\right]. \qquad (1.77)$$

The effective mass $m^*(\hat{k})$ dependence on the wavevector $\mathbf{q}$ direction is expressed in the following way:

$$\frac{1}{m^*(\hat{k})} = \frac{\hat{k}^2_x}{m^*_x} + \frac{\hat{k}^2_y}{m^*_y} + \frac{\hat{k}^2_z}{m^*_z}, \qquad (1.78)$$

with m_x^*, m_y^* and m_z^* effective mass values on the three cartesian axes.
Alternatively, the following formula for the electron energy as a function of k can be employed:

$$\epsilon(\mathbf{k}) = \frac{\sigma}{2} \left(\frac{k_x^2}{m_x^*} + \frac{k_y^2}{m_y^*} + \frac{k_z^2}{m_z^*} \right).$$

(1.79)

2   Froehlich polaron: Feynman diagrams

In order to deeply understand a many-body system such as the Froehlich polaron it is necessary to use a more powerful formalism derived from quantum field theory, we will first breafly describe Green's functions as an effective mean to describe meaningful quantities of our system and we will then shift to the **Matsubara imaginary time** formalism, much more useful in the context of Diagrammatic Monte Carlo. At the end of this journey the connection between the Froehlich polaron and Feynman diagrams will be explicated. In this section of the thesis we will use the convention $\hbar = 1$.

2.1   Green's function formalism

Green functions are useful objects to perturbatively solve systems that are really hard to correctly treat in any other way [21].
If we take for example a time-dependent Schroedinger equation in the following way:

$$[i\partial_t - H_0(\mathbf{r}) - V(\mathbf{r})]\,\psi(\mathbf{r}, t) = 0, \tag{2.1}$$

with the non-interacting diagonizable term H_0 and the perturbation V. It is possible to define the corresponding Green's functions as:

$$[i\partial_t - H_0(\mathbf{r})]\,G_0(\mathbf{r}, \mathbf{r}'; t, t') = \delta(\mathbf{r} - \mathbf{r}')\delta(t - t'),$$
$$[i\partial_t - H_0(\mathbf{r}) - V(\mathbf{r})]\,G(\mathbf{r}, \mathbf{r}'; t, t') = \delta(\mathbf{r} - \mathbf{r}')\delta(t - t'). \tag{2.2}$$

We can define $G_0^{-1}(\mathbf{r}; t')$ and $G^{-1}(\mathbf{r}; t)$ as:

$$G_0^{-1}(\mathbf{r}; t) = i\partial_t - H_0(\mathbf{r}),$$
$$G^{-1}(\mathbf{r}; t) = i\partial_t - H_0(\mathbf{r}) - V(\mathbf{r}), \tag{2.3}$$

The Schroedinger equation can then be recast as

$$\left[G_0^{-1}(\mathbf{r}, t) - V(\mathbf{r})\right]\psi(\mathbf{r}, t) = 0 \tag{2.4}$$

and it is possible to rewrite the system as an integral equation

$$\psi(\mathbf{r}, t) = \psi^0(\mathbf{r}, t) + \int d^3r' \int dt' G_0(\mathbf{r}, \mathbf{r}'; t, t')V(\mathbf{r}')\psi(\mathbf{r}', t') \tag{2.5}$$

which can be solved iteratively. In fact:

$$\psi = \psi^0 + G_0 V \psi^0 + G_0 V G_0 V \psi^0 + G_0 V G_0 V G_0 V \psi^0 + \dots$$
$$= \psi^0 + (G_0 + G_0 V G_0 + G_0 V G_0 V G_0 + \dots)V\psi_0. \tag{2.6}$$

Noting that 2.5 can be also written as:

$$\psi(\mathbf{r}, t) = \psi^0(\mathbf{r}, t) + \int d^3r' \int dt' G(\mathbf{r}, \mathbf{r}'; t, t') V(\mathbf{r}') \psi^0(\mathbf{r}', t'), \tag{2.7}$$

it is possible to identify G with

$$\begin{aligned} G &= G_0 + G_0 V G_0 + G_0 V G_0 V G_0 + \dots \\ &= G_0 + G_0 V (G_0 + G_0 V G_0 + \dots), \end{aligned} \tag{2.8}$$

and the well-known **Dyson equation** is retrieved:

$$G = G_0 + G_0 V G. \tag{2.9}$$

The one described above is the **single particle Green's function**, also called propagator since it "propagates" the wavefunction: in fact, if the wavefunction is known at time t' the wavefunction at a later time can be obtained in the following way:

$$\psi(\mathbf{r}, t) = \int d^3r' \int dt' G(\mathbf{r}t, \mathbf{r}'t') \psi(\mathbf{r}'t'), \tag{2.10}$$

which is a solution to 2.3. The Green's function can also be written as

$$G(\mathbf{r}t, \mathbf{r}'t') = -\theta(t - t')\langle \mathbf{r}|e^{-iH(t-t')}|\mathbf{r}'\rangle, \tag{2.11}$$

and is more precisely known as the **retarded Green's function**.
Focusing now on a many-body system the Green's function is defined as

$$G^R(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') = -i\theta(t - t')\langle[\psi_\sigma(\mathbf{r}, t), \psi_{\sigma'}^\dagger(\mathbf{r}', t')]_{B,F}\rangle, \tag{2.12}$$

where $[,]_{B,F}$ is the commutator $[,]$ for bosons and the anticommutator $\{,\}$ for fermions.
In the case of translation-invariant system (such as lattices) Green's functions can only depend on $\mathbf{r} - \mathbf{r}'$ and it is natural to adopt the usual $\mathbf{k}$ formalism:

$$G^R(\mathbf{r} - \mathbf{r}', \sigma t, \sigma' t') = \frac{1}{V} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} G^R(\mathbf{k}, \sigma t, \sigma' t'), \tag{2.13}$$

$$G^R(\mathbf{k}, \sigma t, \sigma' t') = -i\theta(t - t')\langle[a_{\mathbf{k}\sigma}, a_{\mathbf{k}'\sigma'}^\dagger]_{B,F}\rangle.$$

The goal is now to find explicit expressions for the various Green's function of relevance, in the special case of a free electron the Hamiltonian assumes the simple form

$$H = \sum_{\mathbf{k}\sigma} \epsilon(k) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma}, \tag{2.14}$$

and the time dependence of the creation/annihilation operators is simply defined as

$$c_{\mathbf{k}\sigma}(t) = e^{iHt} c_{\mathbf{k}\sigma} e^{-iHt} = c_{\mathbf{k}\sigma} e^{-i\xi_k t}. \tag{2.15}$$

The retarded Green's function then becomes:

$$G^R(\mathbf{k}\sigma, t - t') = -i\theta(t - t')e^{-i\xi_k(t-t')}. \tag{2.16}$$

2.2 Perturbation theory for Green's functions

Given an Hamiltonian which can be written as a sum of two terms

$$H = H_0 + V, \tag{2.17}$$

where H_0 is diagonizable and V is an interaction term which can be treated perturbatively, we want to obtain a perturbative expansion of the Green's function. The issue with this approach is that in the expression for a generic Green's function it is require to take the expectation value over a ground state ($\langle \cdot \rangle$) which is not known, namely the ground state of the interacting system.

It is thus important to link the ground state of the interacting system to the ground state of the non-interacting one, to this mean we introduce **adiabatic switching on of the interaction** [22].

The core feature of this method is turning a time-independent problem into a time-dependent one by slowly switching on the perturbation, the Hamiltonian is transformed in the following way:

$$H_\alpha(t) = H_0 + V e^{-\alpha|t|}, \qquad \alpha > 0, \tag{2.18}$$

such that the system is non-interacting for $t \to \pm\infty$ and goes back to 2.17 for $t \to \infty$,

$$\lim_{t \to \pm\infty} H_\alpha(t) = H_0, \qquad \lim_{t \to 0} H_\alpha(t) = H. \tag{2.19}$$

In this setup, it is natural to suppose that the ground state of the interacting system $|\Phi_0\rangle$ evolves continuously starting from the ground state of the non-interacting one $|\Psi_0\rangle$.

We introduce the Dirac representation the operators, for the interaction term of the Hamiltonian we have

$$V_I e^{-\alpha|t|} = e^{iH_0 t} V e^{-iH_0 t} e^{-\alpha|t|}, \tag{2.20}$$

the time-evolution operator in the Dirac representation is written as

$$U_{I\alpha}(t, t_0) = \sum_{n_0}^{\infty} \frac{1}{n!} (-i)^n \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n e^{-\alpha(|t_1| + \ldots + |t_n|)} T_t \left[V_I(t_1) \cdots V_I(t_n) \right], \tag{2.21}$$

and defines the evolution of the system in the following way:

$$|\psi_{I\alpha}(t)\rangle = U_{I\alpha}(t, t_0) |\psi_{I\alpha}(t_0)\rangle. \tag{2.22}$$

The equation of motion is obtained as

$$i \frac{d}{dt} |\psi_{I\alpha}(t)\rangle = e^{-\alpha|t|} V_I(t) |\psi_{I\alpha}(t)\rangle, \tag{2.23}$$

from this it follows that

$$i \frac{d}{dt} |\psi_{I\alpha}(t \pm \infty)\rangle = 0, \tag{2.24}$$

which means that the state is time-independent. We can thus take the ground state of the non-interacting system as the ground state for the interacting system when $t \to \pm\infty$, the expression at the two times can differ up to a phase factor:

$$|\psi_{I\alpha}(t \to -\infty)\rangle = |\Psi_0\rangle, \qquad |\psi_{I\alpha}(t \to \infty)\rangle = e^{i\phi}|\Psi_0\rangle. \tag{2.25}$$

It is now possible to link the non-interacting ground state at $t \to -\infty$ with the interacting one at t using the time-evolution operator

$$|\psi_{I\alpha}(t)\rangle = U_I(t, -\infty)|\Psi_0\rangle, \tag{2.26}$$

at $t = 0$ the interaction is completely switched on. Moreover, if the switching on is performed really slowly (adiabatically) it can be excluded that the wavefunction $|\psi_{I\alpha}(t)$ depends on the parameter α and the exact ground state should be computed starting from the non-interacting ground state.
However, since it is required that $\alpha > 0$, there is no certainty that the limit

$$|\Phi_0\rangle = \lim_{\alpha \to 0} |\psi_{I\alpha}(0)\rangle \tag{2.27}$$

exists, and in fact it does not exist.
For this reason the **Gell-Mann-Low theorem** [23] is used, which is a theorem that fixes the eigenstate which evolves from the non-interacting ground state during the adiabatic switching on of the interaction.
The statement of the theorem is the following: given a state

$$\lim_{\alpha \to 0} \frac{U_{I\alpha}(0, -\infty)|\Psi_0\rangle}{\langle\Psi_0|U_{I\alpha,-\infty}|\Psi_0\rangle} = \lim_{\alpha \to 0} \frac{|\psi_{I\alpha}(0)\rangle}{\langle\Psi_0|\psi_{I\alpha}(0)\rangle} \tag{2.28}$$

which exists for every order of the perturbation theory, then this state is an eigenstate $|\Phi_0\rangle$ of the full Hamiltonian H_0.
The importance of this system consists on the fact that it fixes the eigenstate during its evolution starting from the non-interacting ground state, a further important assumption is thus the requirement that no crossings of the states occur during their evolution from the free states. This is not a problem in our case since in the Froehlich Hamiltonian no transitions of this type occur, differently from what happens, for example, in superconductivity.
Since Green's functions are defined as expectation values of Heisenberg operators, these expectation values should be evaluated over the ground state (at $T = 0$), Using Gell-Mann-Low theorem it was demonstrated that the ground state of the interacting system $|\Phi_0\rangle$ can be retrieved from the non interacting one $|\Psi_0\rangle$, the normalized ground state is

$$|\Phi_0\rangle_{norm} = \frac{|\Phi_0\rangle}{(\langle\Phi_0|\Phi_0\rangle)^{1/2}}. \tag{2.29}$$

Given a generic Green's function (in terms of the parameter α)

$$G_\alpha(\mathbf{k}t, \mathbf{k}'t') = -i\langle T_t c_\mathbf{k}(t) c_{\mathbf{k}'}^\dagger(t')\rangle, \tag{2.30}$$

the creation and annihilation operators in the Heisenberg representation can be cast in the interaction one:

$$c_{\alpha\mathbf{k}}(t) = U_{I\alpha}(0,t)c_{I\alpha\mathbf{k}}(t)U_{I\alpha}(t,0). \tag{2.31}$$

The expectation value of the Heisenberg operator can now be computed:

$$
\begin{aligned}
{}_{norm}\langle\Psi_0|c_\mathbf{k}(t)c_{\mathbf{k}'}^\dagger(t')|\Psi_0\rangle_{norm} &= \frac{\langle\Psi_0|c_{\alpha\to0\mathbf{k}}c_{\alpha\to0\mathbf{k}'}^\dagger|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle}\\
&= \lim_{\alpha\to0}\frac{\langle\Psi_0|U_{I\alpha}(+\infty,t)c_{I\mathbf{k}}(t)U_{I\alpha}(t,t')c_{I\mathbf{k}'}^\dagger(t')U_{I\alpha}(t',-\infty)|\Psi_0\rangle}{\langle\Psi_0|U_{I\alpha}(+\infty,-\infty)|\Psi_0\rangle}.
\end{aligned}
\tag{2.32}
$$

Using this formula it is possible to perturbatively treat any Green's function by expanding it in a power series.

Given a generic Heisenberg operator $A(t)$ and the power expansion for the time-evolution operator in 2.21 we may write:

$$
\begin{aligned}
{}_{norm}\langle\Psi_0|A(t)|\Psi_0\rangle_{norm} &= \lim_{\alpha\to0}\frac{1}{\langle\Psi_0|U_{I\alpha}(+\infty,-\infty)|\Psi_0\rangle}\sum_{n=0}^{+\infty}\frac{1}{n!}(-i)^n\cdot\\
&\int_{-\infty}^{+\infty}dt_1\cdots\int_{-\infty}^{+\infty}dt_n e^{-\alpha(|t_1|+\ldots+|t_n|)}\langle\Psi_0|T\left(V_I(t_1)\cdots V_I(t_n)\right)|\Psi_0\rangle.
\end{aligned}
\tag{2.33}
$$

It is then possible to compare 2.33 with 2.21 to obtain the following identity:

$$
\begin{aligned}
\sum_{n=0}^{+\infty}\frac{1}{n!}(-i)^n\int_{-\infty}^{+\infty}dt_1\cdots\int_{-\infty}^{+\infty}dt_n e^{-\alpha(|t_1|+\ldots+|t_n|)}T_t\left(V_I(t_1)\cdots V_I(t_n)A(t)\right) &=\\
=U_{I\alpha}(+\infty,t)A(t)U_{I\alpha}(t,-\infty).
\end{aligned}
\tag{2.34}
$$

The same argument can be made for a single particle Green's function, obtaining a similar expansion:

$$
\begin{aligned}
iG(\mathbf{k},t,t') &= \sum_{n=0}^{+\infty}\frac{1}{n!}(-i)^n\int_{-\infty}^{+\infty}dt_1\cdots\int_{-\infty}^{+\infty}dt_n\cdot\\
&e^{-\alpha(|t_1|+\ldots+|t_n|)}\langle\Psi_0|T_t\left(V_I(t_1)\cdots V_I(t_n)c_\mathbf{k}(t)c_\mathbf{k}^\dagger(t')\right)|\Psi_0\rangle.
\end{aligned}
\tag{2.35}
$$

2.3   Matsubara formalism for imaginary time Green's functions

The usual Green's functions are complex objects that are not fit to be used in the Diagrammatic Monte Carlo method since they provides negative values that cannot be easily sampled from

a distribution without losing precision.

For this reason we now describe the imaginary time formalism for Green's functions, which is also useful to evaluate system that are at non-zero temperature. The relevant substitution is:

$$it \to \tau, \tag{2.36}$$

the retarded Green's function previously defined in 2.13 thus becomes the following Matsubara Green's function:

$$G(\mathbf{k}, \sigma\tau, \sigma'\tau') = -\theta(\tau - \tau')\langle [a_{\mathbf{k}\sigma}, a_{\mathbf{k}'\sigma'}^{\dagger}]_{B,F} \rangle \tag{2.37}$$

where the thermal average $\langle A(\tau)B(\tau') \rangle$ is defined as:

$$\begin{aligned} &\langle \Psi_0 | A(\tau)B(\tau') | \Psi_0 \rangle && T = 0, \\ &\frac{1}{Z} Tr\left[e^{-\beta H} A(\tau)B(\tau') \right] && T > 0. \end{aligned} \tag{2.38}$$

where Ψ_0 is the ground state and Z the partition function.

Given a system with a Hamiltonian characterized by a diagonizable term H_0 and a perturbation $V(\tau)$, we can define the imaginary time Heisenberg picture for an operator A as

$$A(\tau) = e^{-\tau H} A e^{\tau H}, \tag{2.39}$$

and similarly the imaginary time interaction picture as

$$A_I(\tau) = e^{\tau H_0} A e^{-\tau H_0}. \tag{2.40}$$

We can define the imaginary time time-evolution operator in the interaction picture as

$$U_I(\tau, \tau') = e^{\tau H_0} e^{-(\tau - \tau')H} e^{-\tau' H_0}, \tag{2.41}$$

an explicit expression for the interaction picture time-evolution operator in the imaginary time formalism is found in analogy with the real time counterpart

$$\frac{\partial}{\partial \tau} U_I(\tau, \tau') = e^{\tau H_0}(H_0 - H)e^{-(\tau - \tau')H} e^{-\tau' H_0} = -V_I(\tau)U_I(\tau, \tau'), \tag{2.42}$$

which can be solved iteratively:

$$\begin{aligned} U_I(\tau, \tau') &= \sum_{n=0}^{\infty} \frac{1}{n!}(-1)^n \int_{\tau'}^{\tau} d\tau_1 ... \int^{\tau'^{\tau}} d\tau_n T_\tau \left[V_I(\tau_1)...V_I(\tau_n) \right] \\ &= T_\tau \exp\left(-\int_{\tau'}^{\tau} d\tau_1 V_I(\tau_1) \right), \end{aligned} \tag{2.43}$$

where T_τ is the τ ordering operator.

It is also important to stress that the partition function Z is naturally treated in the imaginary time formalism, in fact:

$$e^{-\beta H} = e^{-\beta H_0} U_I(\beta, 0) = e^{-\beta H_0} T_\tau exp\left(-\int_0^\beta d\tau_1 V_I(\tau_1)\right), \qquad (2.44)$$

and we have

$$
\begin{aligned}
\langle T_\tau A(\tau)B(\tau')\rangle &= \frac{1}{Z}Tr\left[e^{-\beta H}T_\tau(A(\tau)B(\tau'))\right] \\
&= \frac{1}{Z}Tr\left[e^{-\beta H_0}U_I(\beta,0)T_\tau(U_I(0,\tau)A_I(\tau)U_I(\tau,\tau')B_I(\tau')U_I(\tau',0))\right] \\
&= \frac{1}{Z}Tr\left[e^{-\beta H_0}T_\tau(U_I(\beta,0)A_I(\tau)B_I(\tau'))\right] \\
&= \frac{\langle T_\tau U_I(\beta,0)A_I(\tau)B_I(\tau')\rangle_0}{\langle U_I(\beta,0)\rangle_0},
\end{aligned}
\qquad (2.45)
$$

where the expectation value is taken on the non-interacting system, in agreement with the previously found result for $T = 0$.

We now go back to the definition of single particle Green's functions in imaginary time formalism, we can define them both in real space and in $\mathbf{k}$ space:

$$
\begin{aligned}
G(\mathbf{r}t, \mathbf{r}'t') &= -\langle T_\tau\left(\psi(\mathbf{r},\tau), \psi^\dagger(\mathbf{r}',\tau')\right)\rangle, \\
G(\mathbf{k}\tau, \mathbf{k}'\tau') &= -\langle T_\tau\left(c_\mathbf{k}(\tau)c_{\mathbf{k}'}^\dagger(\tau')\right)\rangle,
\end{aligned}
\qquad (2.46)
$$

neglecting the spin σ degree of freedom (not relevant for our treatment).

In the non-interacting case the Matsubara single particle Green's functions can be valued in the same way as the retarded Green's functions, in fact:

$$H_0 = \sum_\mathbf{k} \xi_\mathbf{k} c_\mathbf{k}^\dagger c_\mathbf{k}, \qquad (2.47)$$

and the creation/annihilation operators in the Heisenberg picture read

$$c_\mathbf{k}(\tau) = e^{\tau H_0} c_\mathbf{k} e^{-\tau H_0} = c_\mathbf{k} e^{-\xi_\mathbf{k}\tau}, \qquad c_\mathbf{k}^\dagger(\tau) = e^{\tau H_0} c_\mathbf{k}^\dagger e^{-\tau H_0} = c_\mathbf{k}^\dagger e^{\xi_\mathbf{k}\tau}. \qquad (2.48)$$

The non-interacting Matsubara Green's function is then defined as

$$
\begin{aligned}
G_0(\mathbf{k},\tau,\tau') &= -\langle T_\tau\left(c_\mathbf{k}(\tau)c_\mathbf{k}^\dagger(\tau')\right)\rangle \\
&= -\theta(\tau-\tau')\langle c_\mathbf{k}(\tau)c_\mathbf{k}^\dagger(\tau')\rangle - (\pm)_{B,F}\theta(\tau'-\tau)\langle c_\mathbf{k}^\dagger(\tau')c_\mathbf{k}(\tau)\rangle \\
&= -\left[\theta(\tau-\tau')\langle c_\mathbf{k}c_\mathbf{k}^\dagger\rangle - (\pm)_{B,F}\theta(\tau'-\tau)\langle c_\mathbf{k}^\dagger c_\mathbf{k}\rangle\right]e^{-\xi_\mathbf{k}(\tau-\tau')}.
\end{aligned}
\qquad (2.49)
$$

2.4   Equation of motion and Wick's theorem

For a general Green's function the process to find its expression goes through the equation of motion technique, for a general operator in the imaginary time formalism the time derivative is

$$\partial_\tau A(\tau) = \partial(\tau)\left(e^{\tau H}Ae^{-\tau H}\right) = e^{\tau H}[H,A]e^{-\tau H} = [H,A](\tau),\tag{2.50}$$

if we derivate with respect to imaginary time a Matsubara Green's function the following result is obtained:

$$-\frac{\partial}{\partial\tau}\left(T_\tau\langle A(\tau)B(\tau')\rangle\right) = \delta(\tau-\tau')\langle(AB-(\pm)_{B,F}BA)\rangle + \langle T_\tau\left([H,A](\tau)B(\tau')\right)\rangle,\tag{2.51}$$

which for the single particle Green's function in momentum space becomes

$$-\frac{\partial}{\partial\tau}G(\mathbf{k}\tau,\mathbf{k}'\tau') = \delta(\tau-\tau')\delta_{\mathbf{kk}'} + \langle T_\tau\left([H,c_{\mathbf{k}}](\tau)c_{\mathbf{k}'}^\dagger(\tau')\right)\rangle.\tag{2.52}$$

In the case of a Green's function for a non-interacting system G_0 with an Hamiltonian $H_0 = \sum_{\mathbf{kk}'}t_{\mathbf{kk}'}c_{\mathbf{k}}^\dagger c_{\mathbf{k}'}$ this equation simplifies to

$$-\frac{\partial}{\partial\tau}G_0(\mathbf{k}\tau,\mathbf{k}'\tau') = \delta(\tau-\tau')\delta_{\mathbf{kk}'} + \sum_{\mathbf{k}''}t_{\mathbf{kk}''}G_0(\mathbf{k}''\tau,\mathbf{k}'\tau'),\tag{2.53}$$

and the solution is found as previously defined.

We now introduce **Wick's theorem**, an important result which states that higher order (greater than 1) Green's functions for non-interacting systems $G_0^{(n)}$ can be factorized into products of single particle Green's functions G_0.

The n-particle Green's function is defined as:

$$G_0^{(n)}(\nu_1\tau_1,\cdots,\nu_n\tau_n;\nu_1'\tau_1',\cdots\nu_n'\tau_n') = (-1)^n\langle T_\tau\left[c_{I\nu_1}(\tau_1)\cdots c_{I\nu_n}(\tau_n)c_{I\nu_n'}^\dagger(\tau_n')\cdots c_{I\nu_1'}^\dagger(\tau_1')\right]\rangle.\tag{2.54}$$

This expression is quite complicated and difficult to handle, we first define an easier notation to handle all the operators:

$$l_j(\gamma_j) = \begin{cases} c_{I\nu_j}(\tau_j) & \in[0,n], \\ c_{I\nu_{2n+1-j}'}^\dagger(\tau_{2n+1-j}') & \in[n+1,2n], \end{cases}\tag{2.55}$$

and we define the permutation of the $2n$ operators as:

$$P(l_1(\gamma_1)\cdots l_{2n}(\gamma_{2n})) = l_{P1}(\gamma_{P1})\cdots l_{P_{2n}}(\gamma_{P_{2n}}),\tag{2.56}$$

with P_j j-th variable in the permutation P.

The correct order of permutation depends on the the time arguments ordering, for this reason

it is possible to rewrite 2.54 taking into account these considerations:

$$G_0^{(n)}(j_1 \cdots j_{2n}) = (-1)^n \sum_{P \in S_{2n}} (\pm)_{B,F}^P \theta(\gamma_{P_1} - \gamma_{P_2}) \cdots \theta(\gamma_{P_{n-1}} - \gamma_{P_n})$$
$$\times \langle l_{P_1}(\gamma_{P_1}) \cdots l_{P_{2n}}(\gamma_{P_{2n}}) \rangle. \tag{2.57}$$

We can now take the derivative with respect to one of the time variables of $G_0^{(n)}$ and use the equation of motion:

$$\left[-\frac{\partial}{\partial \tau_1} G^{(n)_0} \right]_{2nd-term} = -(-1)^n \left\langle T_\tau \left[H_0, c_{I\nu_1}(\tau_1) \cdots c_{I\nu_n}(\tau_n) c_{I\nu_n}^\dagger(\tau_n') \cdots c_{I\nu_1}^\dagger(\tau_1') \right] \right\rangle \tag{2.58}$$

We know focus on the case where τ_i is next to τ_j', two terms like this are present in 2.57, one where $\tau_i < \tau_j'$ and one where $\tau_i > \tau_j'$ (the permutation order is obviously different).
The Green's function is written as

$$G_0^{(n)} = \left[\cdots \theta(\tau_i - \tau_j') \cdots \right] \langle \cdots c_{I\nu_i}(\tau_i) c_{I\nu_j'}^\dagger(\tau_j') \cdots \rangle$$
$$(\pm)_{B,F} \left[\cdots \theta(\tau_j' - \tau_i) \cdots \right] \langle \cdots c_{I\nu_j'}^\dagger(\tau_j') c_{I\nu_i}(\tau_i) \rangle. \tag{2.59}$$

Differentiating these two terms with respect to τ_i gives two contributions from the heavyside function $\theta(\cdot)$:

$$-\left[\frac{\partial}{\partial \tau_i G_0^{(n)}} \right]_{1st-term} = \left([\cdots] \langle \cdots c_{I\nu_i}(\tau_I) c_{I\nu_j'}^\dagger(\tau_j') \cdots \rangle \right.$$
$$\left. (\mp)_{B,F} [\cdots] \langle \cdots c_{\nu_j'}^\dagger(\tau_j') c_{\nu_i}(\tau_i) \cdots \rangle \right) \delta(\tau_i - \tau_j'). \tag{2.60}$$

We now observe that

$$\left[c_{\nu_i}(\tau_i), c_{\nu_j'}^\dagger(\tau_i) \right]_{B,F} = \delta_{\nu_i, \nu_j'}, \tag{2.61}$$

in the same way, for a couple of terms $c_{\nu_i}(\tau_i)$ and $c_{\nu_j}(\tau_i)$:

$$\left[c_{\nu_i}(\tau_i), c_{\nu_j}(\tau_i) \right]_{B,F} = 0, \tag{2.62}$$

and the same is true for two creation operators.
In this way the number of creation and annihilation operators has been reduced by one each and we have demonstrated that a n order Green's function can written in terms of a $n-1$ Green's functions (which is a sum of different terms). For fermionic Green's functions, the sign of G_0^{n-1} depends on the specific τ_j'. Apart from that, the equation of motion is now:

$$G_{0i}^{-1} G_0^{(n)} = \sum_{j=1}^n \delta_{\nu_i, \nu_j'} \delta(\tau_i - \tau_j')(\pm 1)_{B,F}^x G_0^{(n-1)}(\underbrace{\nu_1 \tau_1, \cdots, \nu_n \tau_n}_{\text{without } i}; \underbrace{\nu_1' \tau_1', \cdots, \nu_n' \tau_n'}_{\text{without } j}). \tag{2.63}$$

The sign $(-1)^x$ for fermions is determined as $(-1)^{n-i+n-j} = (-1)^{-(i+j)} = (-1)^{i+j}$ due to the fact that it is required that the annihilation operator $c_{I\nu_i}(\tau_i)$ is moved i positions to the right while the creation operator $c^\dagger_{\nu'_j}(\tau'_j)$ is moved j positions to the left.

It is now possible to write the explicit decomposition of the n-order Green's function into a sum of n $(n-1)$-order ones:

$$G_0^{(n)} = \sum_{i=0}^{n} (\pm)_{B,F}^{i+j} G_0(\nu_1\tau_1, \nu'_j\tau'_j) G_0^{(n-1)}(\underbrace{\nu_1\tau_1, \cdots, \nu_n\tau_n}_{\text{without } i}; \underbrace{\nu'_1\tau'_1, \cdots, \nu'_n\tau'_n}_{\text{without } j}). \qquad (2.64)$$

It is clear from the obtained result that this method can be naturally generalized to iteratively obtain any non-interacting n-order Green's function as sum of terms only containing products of single particle non-interacting Green's functions, the final result of Wick's theorem is thus obtained:

$$G_0^{(n)}(1, \cdots, n; n', \cdots, 1') = \begin{vmatrix} G_0(1,1') & \cdots & G_0(1,n') \\ \vdots & \ddots & \vdots \\ G_0(n,1') & \cdots & G_0(n,n') \end{vmatrix}_{B,F}, \qquad i \equiv (\nu_i, \tau_i) \quad (2.65)$$

where the result is given as a determinant for fermions and as a permanent for bosons.

2.5   Green's functions for the Froehlich Hamiltonian

We may now consider the Froehlich Hamiltonian previously defined in 1.77 using the new imaginary time formalism, defining the non-interacting Hamiltonian H_0^{cFr} as

$$H_0^{cFr} = \sum_{\mathbf{k}} \frac{\sigma k^2}{2m(\hat{k})} c^\dagger_\mathbf{k} c_\mathbf{k} + \sum_{\mathbf{q}j} \omega_{LO} a^\dagger_{\mathbf{q}j} a_{\mathbf{q}j}. \qquad (2.66)$$

Since our model consists of a single electron (hole) in the minimum (maximum) of the conduction (valence) band interacting with a cloud of phonon, we may consider our system to be at zero temperature and the ground state wavefunction $|\Psi_0\rangle$ to be the vacuum:

$$|\Psi_0\rangle = |0\rangle, \qquad \langle \cdot \rangle = \langle 0| \cdot |0\rangle \qquad (2.67)$$
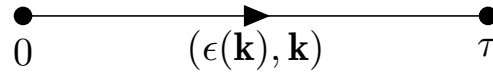
It is then possible to explicitly compute the electron free propagator using 2.49:

$$G_0(\mathbf{k}, \tau, \tau') = -\left[\theta(\tau - \tau')\langle 0|c_\mathbf{k} c^\dagger_\mathbf{k}|0\rangle + \theta(\tau' - \tau)\langle 0|c^\dagger_\mathbf{k} c_\mathbf{k}|0\rangle \right] e^{k^2/2m(\hat{k})(\tau-\tau')}$$
$$= -\theta(\tau - \tau')e^{(k^2/2m(\hat{k}))(\tau-\tau')}, \qquad (2.68)$$

which for $\tau' = 0$ becomes
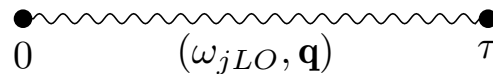
$$G_0(\mathbf{k}, \tau) = -\langle 0|c_\mathbf{k}(\tau)c^\dagger_\mathbf{k}|0\rangle = -e^{-(k^2/2m(\hat{k}))\tau}, \qquad \tau \geq 0. \qquad (2.69)$$

The Matsubara Green's function for a free electron propagator assumes the simple form of an exponential function, which can be easily sampled.

$$-G_0(\mathbf{k}, \tau) = e^{-(k^2/2m(\hat{k}))\tau}$$

Figure 2.1: Feynman diagram of the free electron propagator together with its Green's function.

$$-D_0(\mathbf{q}, \tau) = e^{-\omega_{jLO}\tau}$$

Figure 2.2: Feynman diagram of the free phonon propagator together with its Green's function.

In the same way, it is possible to find an explicit form for the phonon free propagator:

$$D_0(\mathbf{q}j, \tau) = -\langle 0|a_{\mathbf{q}}(\tau)a_{\mathbf{q}}^\dagger|0\rangle = -e^{-\omega_{jLO}\tau}, \qquad \tau \geq 0, \tag{2.70}$$

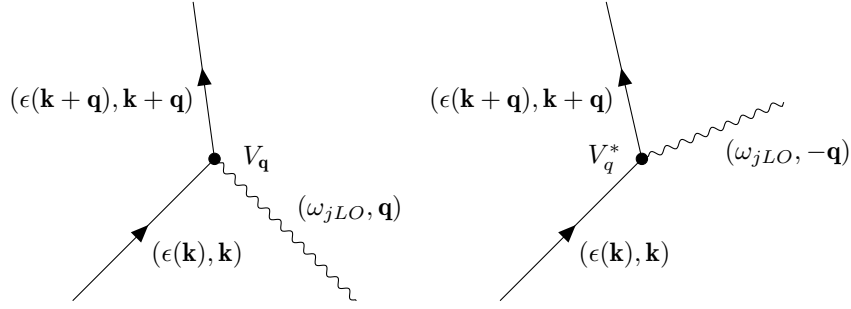which is again a simple exponential function.

Figure 2.3: Feynman diagrams of the electron-phonon interaction.

We now take into consideration the interaction part of the Froehlich Hamiltonian:

$$V = \sum_{\mathbf{k},\mathbf{q}j} c_{\mathbf{k}+\mathbf{q}}^{\dagger} c_{\mathbf{k}} \left[V_{\mathbf{q}}^{cFr} a_{\mathbf{q}} + V_{\mathbf{q}}^{*cFr} a_{-\mathbf{q}}^{\dagger} \right], \tag{2.71}$$

The basic Feynman diagram that describes this interaction is an interaction between two free electron propagators (one with momentum $\mathbf{k}$ gets annihilated and one with momentum $\mathbf{k} + \mathbf{q}$ gets created) and a free phonon propagator which can either be annihilated (momentum $\mathbf{q}$) or created (momentum $-\mathbf{q}$), the strength of the interaction at the vertex is given by $|V_{\mathbf{q}}|$ (2.3). We have now obtained the fundamental building blocks which can be used to obtain the total Green's function for the Froehlich polaron $G(\mathbf{k}, \tau)$, a generic one-electron Matsubara Green's function for the polaron is written as [24]

$$G(\mathbf{k}, \tau - \tau') = -\langle c_{\mathbf{k}}(\tau) c_{\mathbf{k}}^{\dagger}(\tau') \rangle = -\langle 0 | c_{\mathbf{k}}(\tau) c_{\mathbf{k}}^{\dagger}(\tau') | 0 \rangle. \tag{2.72}$$

The creation and annihilation operators here defined do not have the simple form described above for the free propagators, it is thus necessary to use the interaction picture and thus transform the operators:

$$c_{I\mathbf{k}}(\tau) = e^{\tau H_0^{cFr}} c_{\mathbf{k}} e^{-\tau H_0^{cFr}}, \qquad c_{I\mathbf{k}}^{\dagger}(\tau) = e^{\tau H_0^{cFr}} c_{\mathbf{k}}^{\dagger} e^{-\tau H_0^{cFr}}. \tag{2.73}$$
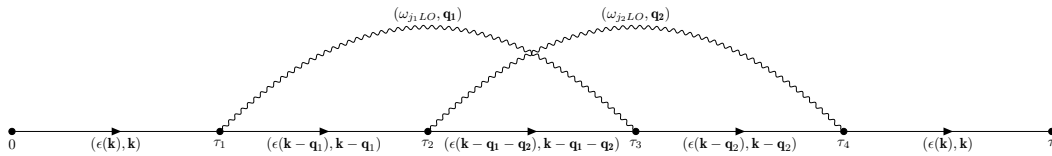
Figure 2.4: Order 4 diagram.
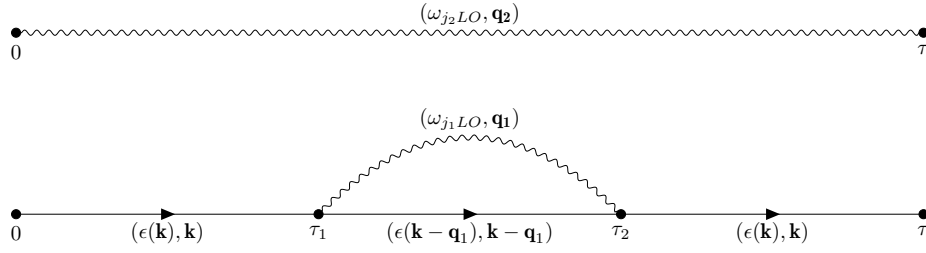
Figure 2.5: Disconnected order 4 diagram.

In the interaction picture the Green's function for the Froehlich polaron assumes the form [25]:

$$G(\mathbf{k}, \tau) = - \left\langle 0 \left| T_\tau \left[c_{I\mathbf{k}}(\tau) c_{\mathbf{k}}^\dagger \exp\left(-\int_0^{+\infty} V_I(\tau') d\tau' \right) \right] \right| 0 \right\rangle_{\text{conn}}, \qquad (2.74)$$

where the expectation value is restricted to connected diagrams, which means that no integral over the internal variables $d\tau_i'$ can be represented by an external factor (the difference between a connected and disconnected diagram is easily recognazible in 2.4 and 2.5).
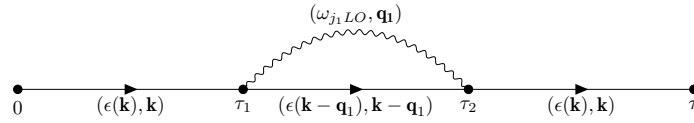
Figure 2.6: Order 2 diagram.
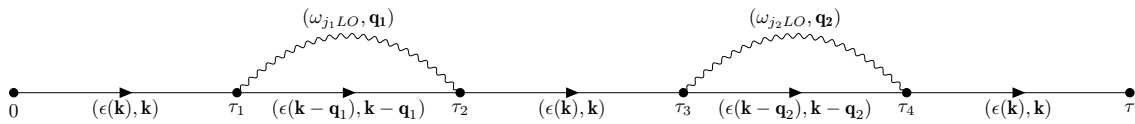
Figure 2.7: Order 4 diagram.
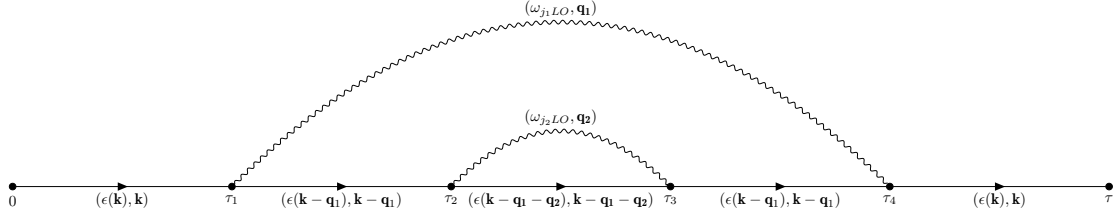
Figure 2.8: Order 4 diagram.

Using Wick's theorem, the product of n chronologically ordered operators $\{\tau_1, \tau_2, ..., \tau_n\}$ can be evaluated as the sum of products of pairs of operators, such that it is possible to expand the Green's function as an infinite series of terms of the following form:

$$G(\mathbf{k}, \tau) = - \sum_{n=0,2,4...}^{+\infty} \sum_{\xi_n} \int d\tau_1 \cdots \int d\tau_n D_n^{\xi_n}(\mathbf{k}, \tau; x_1, ..., x_n), \qquad (2.75)$$

where n indexes the rank of the term (only terms with an even number of internal variables are allowed), ξ_n indexes the topology of the term (which depends on how the phonon lines are arranged, see the difference between 2.7, 2.4 and 2.8), x_i are the internal variables which index the imaginary times at which a vertex $V^{cFr}(qj)$ is present (and a phonon is created or annihilated), the index j of the phonon mode which interacts with the electron propagators and the momentum q of the created/annihilated phonon line.
As an example, the 0-order term is just the free electron propagator $G_0(\mathbf{k}, \tau) = -e^{-\epsilon(\mathbf{k})\tau}$, the 2-order term (with internal variables (τ_1, j_1) and (τ_2, j_2)) represents a diagram with 3 free electron propagators, 1 free phonon propagator and 2 vertices.
$D_2(\mathbf{k}, \tau; x_1, x_2,)$ can thus be represented using Green's function as

$$\begin{aligned} D_2(\mathbf{k}, \tau; x_1, x_2,) = &|V^{cFr}(qj)|^2 D_0(\mathbf{q}j, \tau_2 - \tau_1) G_0(\mathbf{k}, \tau - \tau_2) \times \\ &G_0(\mathbf{k} - \mathbf{q}, \tau_2 - \tau_1) G_0(\mathbf{k}, \tau_1), \end{aligned} \qquad (2.76)$$

which translates to

$$D_2(\mathbf{k}, \tau; x_1, x_2,) = |V^{cFr}(qj)|^2 e^{-\omega_{jLO}(\tau_2 - \tau_1)} e^{-\epsilon(\mathbf{k})(\tau - \tau_2)} e^{-\epsilon(\mathbf{k} - \mathbf{q})(\tau_2 - \tau_1)} e^{-\epsilon(\mathbf{k})\tau_1}. \qquad (2.77)$$

It should be noted that in our computation instead of 2.75 the result without the minus sign will be taken: this is just a convention used in order to obtain a definite positive distribution for the Monte Carlo sampling and it does not affect the physical significance of the calculated Green's functions.
Having seen the one-electron Green's function we now introduce another related Green's function: the one-electron N-phonons Green's function $G^N(\mathbf{k}, \sum_{j=1}^{N} \tilde{\mathbf{q}}_j, \tau)$. Differently from

the one-elecron Green's function, the one-electron N-phonons Green's function is a more complex object consisting of N phonons and one electron, and better describes the polaron physics, expecially if the strength of the interaction is intense.

Its expression, using Wick's theorem, can be written as [26]:

$$G^N(\mathbf{k}, \sum_{j=1}^N \tilde{\mathbf{q}}_j, \tau) = \langle 0|a_{\tilde{\mathbf{q}}_1}(\tau)...a_{\tilde{\mathbf{q}}_N}(\tau)c_{\mathbf{p}}(\tau)c_{\mathbf{p}}^\dagger a_{\tilde{\mathbf{q}}_N}...a_{\tilde{\mathbf{q}}_1}|0\rangle\,, \qquad (2.78)$$

where $\mathbf{p} = \mathbf{k} - \sum_{j=1}^N \tilde{\mathbf{q}}_j$.

It is worth noting that $G^N(\mathbf{k}, \sum_{j=1}^N \tilde{\mathbf{q}}, \tau)$ will only take into account connected one-electron N-phonons Green's functions since in the other cases the external phonons can just be neglected. The related expansion can be given as:

$$G^N(\mathbf{k}, \sum_{j=1}^N \tilde{\mathbf{q}}_j, \tau) = \sum_{n=0,2,4...}^{+\infty} \sum_{\xi_n} \int d\tilde{\mathbf{q}}_1 \cdots \int d\tilde{\mathbf{q}}_N \cdot$$
$$\cdot \int d\tau_1 \cdots \int d\tau_n D_n^{\xi_n}(\mathbf{k}, \sum_j \tilde{\mathbf{q}}_j, \tau; x_1, ..., x_n), \qquad (2.79)$$

where the minus sign has been dropped.

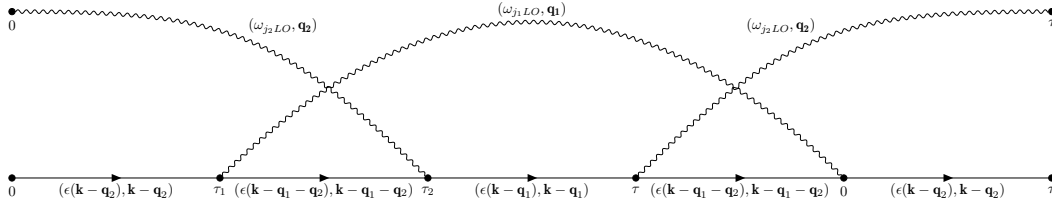Figure 2.9: Order 2 diagram with one external phonon.

It is thus useful to consider the full function $P(\mathbf{k}, \tau)$ defined as follows:

$$P(\mathbf{k}, \tau) = G(\mathbf{k}, \tau) + \sum_{N=1}^{+\infty} G^N(\mathbf{k}, \sum_{j=1}^N \tilde{\mathbf{q}}_j, \tau), \qquad (2.80)$$

which more accurately describes our polaron model, including both the "free" polaron and its interaction with the phonon cloud.

3   The Diagrammatic Monte Carlo method

3.1   The Monte Carlo sampling method

The Monte Carlo method is not a specific technique, but rather a wide set of similar methods which employs probability to solve problems that are otherwise too complex for an analytic solution and too resource demanding to solve with a numerical method.

The basic idea behind Monte Carlo is to use a statistical approach in the resolution of difficult integral and differential equations [27] by the means of a precisely defined set of rules and a random number generator.

This method is an *iterative stochastic procedure*, or in layman terms, a procedure that is iterative: it needs to be applied many times in order to produce an extremely large number of measurements from which it is possible to build an estimate for a determined quantity using the **central limit theorem** and the law of large numbers, and stochastic: it uses random numbers to obtain all sorts of distributions, usually through a precisely defined set of rules using a **Markov chain**.

It is in fact possible, using a Monte Carlo method, to estimate the ground state of the time-independent Schroedinger equation [4]

$$-\nabla^2 \psi(x,y,z) = [E - U(x,y,z)]\,\psi(x,y,z) \tag{3.1}$$

using the following ansatz for the wavefunction:

$$u(x,y,z,\tau) = \psi(x,y,z)e^{-E\tau}, \tag{3.2}$$

thus introducing a fictitious imaginary time-dependence. In this way $u(x,y,z,t)$ follows the diffusion equation

$$\frac{\partial u}{\partial \tau} = \nabla^2 u - Uu, \tag{3.3}$$

which can be framed in a Monte Carlo representation as a set of weighted particles which independently perform a random walk with an exponential decay in imaginary time with a rate governed by the energy eigenvalue E, together with a particle distribution which can be used to determine an estimate for the wavefunction $\psi(x,y,z)$.

Note that the transformation performed on the wavefunction (first done by Fermi) is exactly the already seen transformation used for the Matsubara Green's functions 2.36, which turns the standard time-dependent Schroedinger equation

$$i\frac{\partial \psi}{\partial t} = -\nabla^2 \psi + U\psi \tag{3.4}$$

into

$$\frac{\partial \psi}{\partial \tau} = \nabla^2 \psi - U\psi. \tag{3.5}$$

This observation already stresses the importance of imaginary time in our computation. We list three main types of Monte Carlo simulations [8]:

- **Direct Monte Carlo**, where the generation of random numbers directly models a physical system (usually with a random walk) without directly defining the complexities which characterizes it. An example is the aforementioned model to solve the ground state of the Schroedinger equation.

- **Monte Carlo integration**, a method that is specifically used to compute hard integrals using random numbers.

- **Markov chain Monte Carlo**, which generates the distribution of a system using a Markov chain. This method is used to study the properties of classical and quantum systems.

3.2 Direct Monte Carlo

In direct Monte Carlo the expectation value $\langle I \rangle$ of a variable is estimated, which means that we compute its mean. In order to do so the deterministic problem must be recast in a probabilistic form. $\langle I \rangle$ is a number, in particular it can be seen as the result of an integral.

Given X a random variable defined on a set Ω, we can define $\langle I \rangle$ as the expectation value $E(X)$ of the random variable. In statistics, the expectation value $E(X)$ and the variance $\sigma^2(X)$ have the following definitions:

$$\langle I \rangle = E(X) = \int_\Omega dp X, \qquad \sigma_I^2 = \sigma^2(X) = \int_\Omega dp (X - E(X))^2, \tag{3.6}$$

where p is the probability measure [28].

An approximate estimate for the value I is obtained by producing an independent sequence of random event ω_i according to the probability law p with value

$$\bar{I}_N = E(X_N) = \frac{1}{N} \sum_{i=1}^{N} X(\omega_i), \tag{3.7}$$

with $\bar{I}_N$ the arithmetic mean of N random events.

Given the fact that I_N is just an estimate of the expectation value $\langle I \rangle$, it is important to estimate its deviation from the exact value, to this reason we introduce the **Chebyshev's inequality** [29], which states that no more than $1/k^2$ of the distribution values can be more than k standard deviations from the mean value:

$$P(|x - \langle x \rangle| > k\sigma) \leq \frac{1}{k^2}. \tag{3.8}$$

Chebyshev's inequality makes no assumptions on the distribution and is thus very general, yet it provides an upper bound for the probability to find random values far from the mean.

In the case of N independent random variable for which

$$P(x_1, x_2, ..., x_N) = P(x_1)P(x_2)...P(x_N) \tag{3.9}$$

holds true, it is possible to consider a new random variable z which is the sum of the N original random variables. If the random variables X_i have generic distribution functions, then the distribution function of the sum has a complicated form, the important consideration is that, under very broad assumptions, it is possible to obtain an asymptotically exact form for the distribution function of z when the number N of independent variables becomes very large. Given

$$\bar{x}_N = \frac{1}{N} \sum_{i=0}^{N} x_i, \tag{3.10}$$

both the expectation value and the variance of $\bar{x}_N$ can be easily computed, since all the N terms of the sum give an identical contribution equal to $\langle x \rangle$, resulting in

$$\langle \bar{x}_N \rangle = \langle x \rangle. \tag{3.11}$$

This results holds also in the case where the N random variables x_i are not independent. The variance of $\bar{x}_N$ is easily obtained in the following way:

$$\sigma_{\bar{x}_N}^2 = \langle \bar{x}_N^2 \rangle - \langle \bar{x}_N \rangle^2 = \frac{\sigma^2}{N}. \tag{3.12}$$

We thus have for $N \to \infty$ the random variable $\bar{x}_N$ which has a very narrow distribution centered about the true expectation value of the random variable X (noting that $\sigma_{\bar{x}_N}^2 \to 0$). This result is called the **weak law of large numbers**.

With the **central limit theorem** instead we obtain the asymptotic probability distribution of the sum z of a large number of random variables which are independent and equally distributed.

Let us define the random variable Y as

$$Y = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} y_i = \sqrt{N}(\bar{x}_N - \langle x \rangle), \tag{3.13}$$

which has the expectation value

$$\langle Y \rangle = 0. \tag{3.14}$$

The characteristic function of Y is given by

$$\phi_Y(t) = \left[\phi_y \left(\frac{t}{\sqrt{N}} \right) \right]^N, \tag{3.15}$$

assuming that all the y_i are independent and identically distributed.

Expanding the characteristic function up to second order and taking the limit for $N \to \infty$ we get:

$$\lim_{N \to \infty} \phi_Y(t) = \exp\left(-\frac{\sigma^2 t}{2}\right),$$

(3.16)

which is the characteristic function of a gaussian random variable:

$$P(Y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{Y^2}{2\sigma^2}\right),$$

(3.17)

and we obtain for $\bar{x}_N = \langle x \rangle + Y/\sqrt{N}$ the same gaussian distribution with mean $\langle x \rangle$ and variance σ^2/N.

Having now stated these necessary elements from probability and statistics it is easy to see how a direct Monte Carlo method works: we directly employ random numbers to compute quantities, exploiting the fact that given N iterations and $N \to \infty$, we will obtain a convergent solution for the modelled problem. The big issue is, of course, representing the investigated phenomenon in a way that makes it possible to use random numbers.

3.3   Monte Carlo integration

The Monte Carlo integration method is a technique specifically employed to compute integrals. Consider a generic integral of a smooth generic function $f(\mathbf{x})$ of vector $\mathbf{x}$ and d components:

$$I = \int f(\mathbf{x}) d\mathbf{x},$$

(3.18)

in Monte Carlo integration we recast this integral in the following way using a probability distribution $p(x)$, ($\int p(x) = 1$):

$$\left\langle \frac{f(\mathbf{x})}{p(\mathbf{x})} \right\rangle = \int \frac{f(\mathbf{x})}{p(\mathbf{x})} p(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}) d\mathbf{x}.$$

(3.19)

The integral recast in this form is the expectation value of the function $f(\mathbf{x})$ divided by the probability distribution $p(\mathbf{x})$. The central limit theorem then implies that it is possible to estimate the integral I (deterministic) as the average value of $f(x)$ over a large number of sampling of the random variable $\mathbf{x}_i$ with distribution $p(\mathbf{x})$:

$$I \approx I_N = \frac{1}{N} \sum_{i=1}^{N} \frac{f(\mathbf{x}_i)}{p(\mathbf{x}_i)},$$

(3.20)

where the random variables $\mathbf{x}_i$ are sampled according to $p(\mathbf{x}_i)$.

For large N I_N is normally distributed with mean equal to I and variance σ^2/N computed as

$$\sigma^2 = \left\langle \left(\frac{f(\mathbf{x})}{p(\mathbf{x})}\right)^2 \right\rangle - \left\langle \frac{f(\mathbf{x})}{p(\mathbf{x})} \right\rangle^2,$$

(3.21)

thus for $N \to +\infty$ I_N tends to the deterministic value I.

It is now relevant to address the main issue with this method: generating configurations $\mathbf{x}_i$ that are distributed according to $p(\mathbf{x})$ and then compute $f(\mathbf{x}_i/p(\mathbf{x}_i))$. When it is possible to directly generate samples from $p(\mathbf{x})$ we are employing **direct MC sampling**, this is possible if we are using a uniform or exponential distribution (or other similar simple distributions). The samples generated with this method are independent, but we are limited in the type of distribution that can be used, which could make the convergence to the exact value extremely slow.
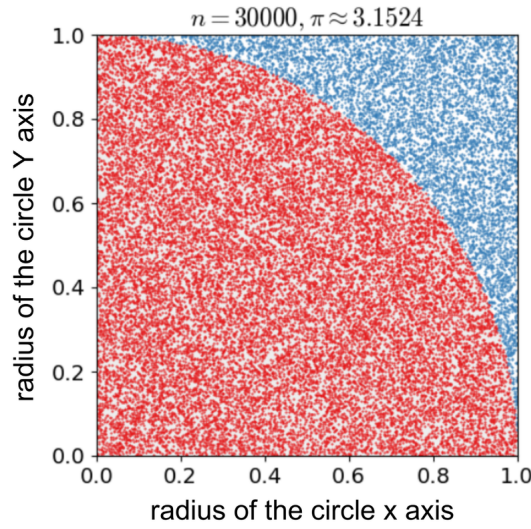
Figure 3.1: Graphical representation of π estimation using the simple direct Monte Carlo integration method.

The simplest way to show the direct sampling method for integrals is through the computation of π: suppose that we have a square of side 1 and, inside it, the quadrant of a circle (of radius 1). It is known from high school geometry that:

$$\frac{\pi}{4} = \frac{1}{4}\frac{\pi(1)^2}{(1)^2} = \frac{1}{4}\frac{A_{\text{quadrant}}}{A_{\text{square}}}, \tag{3.22}$$

from which it follows that we can compute the value of π from the ratio between the two areas. We now express the areas in terms of integrals

$$\frac{\pi}{4} = \frac{\int_{\text{circle}} dxdy}{\int_{\text{square}} dxdy} = \int_{\text{square}} f(x,y)p(x,y)dxdy, \tag{3.23}$$

with

$$p(x,y) = \frac{1}{\int_{\text{square}} dxdy} \tag{3.24}$$

and $f(x, y)$ defined in the following way:

$$f(x, y) = \begin{cases} 1 & \text{, if } \sqrt{x^2 + y^2} < 1 \\ 0 & \text{, otherwise}. \end{cases} \tag{3.25}$$

Since $p(x, y)$ is non-negative and normalized, it can be treated as a probability distribution. We can thus say that

$$\frac{\pi}{4} \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i, y_i), \tag{3.26}$$

the obtained result can be interpreted visually considering that we are filling the total area of the square with randomly distributed dots and counting how many of them end inside the circle compared to the total.

This simple method can be used (at least in theory) to compute every possible integral, nevertheless using a uniform distribution may result impractical when estimating functions that have sharp peaks: in this cases most of the guesses made end up "outside" the region of interest and do not contribute to the estimation of the integral. Of course, this problem become more and more relevant the more dimensions we have.

In this cases, it is appropriate to use as a probability function that has a similar shape with respect to the target function we want to integrate, consider the following one-dimensional case:

$$I = \int_a^b f(x)dx, \tag{3.27}$$

where $f(x)$ has a sharp peak in a limited interval between a and b and is close to 0 everywhere else. Operating in the same way as before we obtain the following expression:

$$I = \int_a^b \frac{f(x)}{p(x)} p(x)dx, \tag{3.28}$$

and we can approximate the integral as

$$I \approx I_N = \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)}, \tag{3.29}$$

while the variance of our estimation is found as

$$s^2 = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{f(x_i)}{p(x_i)} \right]^2 - \left[\frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)} \right]^2, \tag{3.30}$$

analyzing the expression for the variance estimator it becomes clear that it is minimized if $p(x)$ is chosen to be as close as possible to the target function $f(x)$, with the best result $s^2 = 0$ obtained for $p(x) = cf(x)$ (which would mean that it is possible to analytically integrate $f(x)$).

This means that with a careful choice of the probability distribution it is possible to minimize the statistical fluctuations of our estimate and thus obtain the same accuracy with a smaller number of samples N.

Since multiple algorithms have been studied in order to build pseudonumber generators in the interval $(0, 1)$ that are efficient, provide long periods for sequences of values and have uniform distribution in n-dimensional spaces (for example the Mersenne-Twister algorithm, used in this thesis [30]), it is important to find ways to sample from other more complex (continuous) distributions $p(x)$, the easiest way to do so is through **CDF inversion** (cumulative density function inversion), which, as the name suggests, can be used when the analytical form of the cumulative distribution $P(x)$ is known.

Choosing for example the exponential distribution

$$\exp(x; \lambda) = \lambda e^{-\lambda x} \qquad \text{for } x > 0, \tag{3.31}$$

for which we can obtain the expression for the cumulative density function easily:

$$P(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}. \tag{3.32}$$

We now solve for x to obtain

$$x = -\frac{1}{\lambda} \log(1 - P(x)). \tag{3.33}$$

Since by definition the image of $P(x)$ is $[0, 1]$, we can substitute it in the equation with the standard uniform random variable defined in $[0, 1]$ r:

$$x = -\frac{1}{\lambda} \log(1 - r), \tag{3.34}$$

and we obtain a random variable x that follows the exponential distribution $p(x; \lambda)$.

3.4   Markov Chain Monte Carlo

The procedure illustrated in the previous section only works when an analytical expression for the cumulative distribution exists, this is not true in general and a different method is thus needed for a general distribution $p(x)$. To this scope **Markov chains** are used.

Let us define the random vector $\mathbf{X}$ as a sequence of n random samples that are drawn one after the other:

$$\mathbf{X} = (X_1, X_2, ..., X_n), \tag{3.35}$$

we call this vector an uncorrelated chains if each one of the variables x_i are independent and thus uncorrelated, the probability of obtaining the chain $\mathbf{X}$ is then given by the product of the individual elements of the chain:

$$P_n(\mathbf{X}) = P_1(X_1)P_2(X_2)\cdots P_n(X_n). \tag{3.36}$$

In a Markov chain this is not true and the random samples are correlated: a sequence of samples is said to be Markovian if the conditional probability of each sample of the sequence satisfies

$$P(X_i|X_{i-1}...X_1) = P(X_i|X_{i-1}).\qquad(3.37)$$

This means that the probability of X_i only depends on the previous element X_{i-1} (and nothing else), with the exception of the first element of the chain. This means that the order of the variables is important in the determination of the probability of the full random vector:

$$P_n(X_1, X_2, ..., X_n) = P_1(X_1)P(X_2|X_1)\cdots P(X_n|X_{n-1})\qquad(3.38)$$

where $P(X_i|X_{i-1})$ is the transition probability from X_{i-1} to X_i. We define $p_i^{(1)}$ ($\sum_i p_i^{(1)} = 1$) the initial probability and $P(X_i|X_j) = P_{ij}$ the transition probability from j to i. The transition probabilities are normalized for each j such that $\sum_i P_{ij} = 1$, in this way for every k step $\sum_k p_i^{(k)} = 1$ is obtained (note that, in continuous distribution, the sum is replaced by an integral).

Apart from the (obvious) normalization constraint, the single probabilities $p_i^{(k)}$ at the step k (which are steps in time) are in general different from the probabilities $p_i^{(k-1)}$, the important mathematical result that is obtained under very general assumptions is that, after a necessary relaxation time that is different from system to system, the distribution relaxes to a stationary state p_i which satisfies the relation

$$p_i = \sum_j P_{ij}p_j\qquad(3.39)$$

for every k step in the simulation.

The conditions required for this to happen are:

- **connectedness**: every configuration of the distribution that we want to obtain must be accessible from any other configuration of the system in a finite number of steps.

- **no periodicity**: which means that, after having visited a specific configuration, it must be false that the same configuration will be accessible again only after nk steps with n positive integer and k fixed.

A Markov chain that satisfies these two conditions is called **ergodic** [4].

Since the objective is to obtain a target stationary distribution p_i, we need a mean to do so starting from a more generic distribution (such as the uniform distribution in $[0, 1]$), which means that a specific expression for the transition probabilities P_{ij} is needed. The solution proposed by Metropolis [31] is to require the following relation for the transition probability P_{ij} to be true:

$$P_{ij}p_j = P_{ji}p_i,\qquad(3.40)$$

which is the **detailed balance condition**. This means that the probability of transitioning from j to i is the same as the probability of transitioning from i to j. This relation can be demonstrated using the **Master equation**:

$$p_i^{(k)} - p_i^{(k-1)} = -\sum_j P_{ji} p_i^{(k-1)} + \sum_j P_{ij} p_j^{(k-1)}, \tag{3.41}$$

the stationary condition is obtained for

$$p_i^{(k)} = p_i^{(k-1)} = p_i, \tag{3.42}$$

From this it follows that

$$\sum_j P_{ji} p_i = \sum_j P_{ij} p_j, \tag{3.43}$$

for every i, j pair.
An obvious solution to this equation is found if 3.40 holds.
We now recast the transition probability matrix P_{ij} in the form

$$P_{ij} = T_{ij} A_{ij}, \tag{3.44}$$

where T_{ij} is the symmetric matrix ($T_{ij} = T_{ji}$) of the trial transition proposal probabilities with the additional properties $0 \leq T_{ij} \leq 1$ and $\Sigma_i T_{ij} = 1$, while A_{ij} are the elements of the acceptance matrix, whose form is

$$A_{ij} = \min \left\{ 1, \frac{p_i}{p_j} \right\} \qquad i \neq j. \tag{3.45}$$

Using this configuration we can show that the condition of detailed balance is obeyed by assuming $p_i < p_j$ (we are not posing any constraint on the system by doing this since the choice of the indexes is arbitrary) and observing that

$$P_{ij} = T_{ij} \frac{p_i}{p_j} = T_{ji} \frac{p_i}{p_j} = T_{ji} \cdot 1 \cdot \frac{p_i}{p_j} = P_{ji} \frac{p_i}{p_j}, \tag{3.46}$$

from which 3.40 follows.
The complete algorithm, known as the **Metropolis algorithm**, proposes a new i state given the present j with trial transition probability T_{ij}, then uses the A_{ij} elements of the acceptance matrix to accept i or reject it (and thus keep j). Whether the new i state is accepted or rejected depends on the uniform random variable defined in $[0, 1]$ r: if $r < A_{ij}$ i is accepted, otherwise it is rejected.

Algorithm 3 Metropolis algorithm.

Input: Proposal probability T, limiting distribution p, configuration j.

  Sample an i from T_{ij} ;

  Generate a uniform random number $\zeta \in [0, 1]$;

  if $\zeta \leq p_i/p_j$ **then**

    $j \leftarrow i$;

  end if

  return j.

Figure 3.2: Schematic representation of the Metropolis algorithm procedure [4].

It is possible to generalize the algorithm to one that does not have a symmetric trial transition probability matrix T_{ij}.

Given a generic acceptance ratio R_{ij} such that $A_{ij} = \min\{1, R_{ij}\}$ defined as:

$$R_{ij} = \frac{S_{ij}}{1 + \frac{T_{ij}p_j}{T_{ji}p_i}}, \tag{3.47}$$

with S_{ij} non-negative symmetric matrix.

If S_{ij} is chosen to be:

$$S_{ij} = \begin{cases} 1 + \frac{T_{ji}p_i}{T_{ij}p_j}, & T_{ij}p_j \geq T_{ji}p_i, \\ 1 + \frac{T_{ij}p_j}{T_{ji}p_i}, & T_{ji}p_1 \geq T_{ij}p_j, \end{cases} \tag{3.48}$$

the **Metropolis-Hasting algorithm** is obtained with acceptance ratio R_{ij}:

$$R_{ij} = \frac{T_{ji}p_i}{T_{ij}p_j}. \tag{3.49}$$

In the special case where we want to obtain a random variable X distributed as a generic continuous function $f(x)$ (not necessarily normalized) using as a prior distribution a known distribution $h(x)$, we may identify the trial transition probability from a state x to a state x' with

$$T_{x'x} = \pi(x')dx \tag{3.50}$$

independent on the current state x, the acceptance ratio $R_{x'x}$ then becomes

$$R_{x'x} = \frac{T_{xx'}p_{x'}}{T_{x'x}p_x} = \frac{\pi(x)dx\,f(x')dx}{\pi(x')dx\,f(x)dx} = \frac{\pi(x)f(x')}{\pi(x')f(x)}, \tag{3.51}$$

with acceptance probability $A_{x'x}$:

$$A_{x'x} = \min\left\{1, \frac{\pi(x)f(x')}{\pi(x')f(x)}\right\}. \tag{3.52}$$

Using the above formulation (easily generalizable at least in theory to N dimensions) it is possible to sample random variable from any generic continuous unnormalized function. The main drawbacks of this method are the fact that a relaxation to the stationary distribution is necessary (and the number of steps required is not known a priori) and that the obtained samples are correlated.

3.5 Diagrammatic Quantum Monte Carlo

In Chapter 2 we have seen that the polaron model can be solved by perturbatively expanding the related Green's function (which is a solution to the equation of motion of the system) in the imaginary time formalism. Since the Green's function in this formalism can be cast in a form that makes it real and non-negative, it can be sampled as if it were a distribution using the methods provided by Markov Chain Monte Carlo.

For this reason, we need a method to ergodically represent all the possible configurations of the polaron system. This approach can be used in many different condensed matter physics where a similar expansion is possible, provided that the weight of negative diagrams (which in our specific case do not appear, but are generally present in fermionic systems) does not inficiate the errors of the estimated quantity [32].

Given a function represented as 2.75 (without the minus sign) it is possible to sample it as a distribution $Q(\{y\})$ which depends on a set of variable y:

$$Q(\{y\}) = \sum_{n=0}^{+\infty} \sum_{\xi_n} \int dx_1 \cdots \int dx_n D_n(\xi_n, y, x_1, ..., x_n), \qquad (3.53)$$

with $\{y\}$ external variables and $\{x_i\}$ internal variables (to be integrated out).

Using the Metropolis-Hastings algorithm it is possible to sample variables from this distribution and in the limit of a large number of samples retrieve the full function up to a normalization constant C with no approximations:

$$Q_i(\{y\}) = C \ G(\mathbf{k}, \tau). \qquad (3.54)$$

The fundamental difference with respect to an usual Monte Carlo integration is the fact that the parameter space of internal variables varies between different iterations of the simulation: ways to model transitions between these different states together with the correct trial transition probability are required.

In order to do so, a number of elementary updates are defined such that there are at least enough to satisfy the ergodicity requirement [4] for the Markov chain which relaxes to the target distribution (the system under study). The Metropolis-Hastings algorithm in DMC uses as acceptance probabilities for transitions between two different states characterized by the variables $i = \{y_i, \xi_i, \mathbf{x}_i\}$ and $i = \{y_j, \xi_j, \mathbf{x}_j\}$ respectively the relative weight of the two diagrams.

In fact, given the two diagrams $D_{n_i}(i)$ and $D_{n_j}(j)$ (which may be different in the value of one or more variables, whether internal or external, and more importantly in the *number* of internal variables if the rank of the two diagrams is different), the acceptance ratio R_{ij} is proportional to:

$$R_{ij} \propto \frac{D_{n_i}(i)}{D_{n_j}(j)} = \frac{C\ D^*_{n_i}(i)}{C\ D^*_{n_j}(j)} = \frac{D^*_{n_i}(i)}{D^*_{n_j}(j)}, \tag{3.55}$$

where the $D^*_{n_k}(k)$ are the properly normalized weights, different from the weights of the generated distribution up to a normalization constant [5].

In the case of negative diagrams, the acceptance ratio is performed on the absolute values of the weights. Operating in this way, however, we expose ourselves to the **negative sign problem**: when we want to compute an observable O from the Monte Carlo simulation, whether it be the (normalized) distribution or any other quantity of interest, we must take into account the negative weights:

$$O = O_N = \left(\frac{1}{N} \sum_{i=1}^{N} |O_i(D(i))| \right) \left(\frac{N^+ - N^-}{N} \right)^{-1} = \frac{1}{N^+ - N^-} \sum_{i=1}^{N} |O_i(D(i))|, \tag{3.56}$$

and the variance of the Monte Carlo estimate diverges for $N^+ - N^- \approx 0$.

Two basic classes of elementary updates can be recognized [33][24]:

- **class I updates**: they change one or more variables (internal or external), the new variables $\mathbf{x}$ are proposed using a suitable distribution $\pi(\mathbf{x})$ (some distribution might be more suited than others depending on the specific system), and the acceptance ratio is computed using the standard Metropolis-Hastings algorithm (3.49 and 3.51 depending on the type of variable and prior distribution)

- **class II updates**: in this class of updates the order of the diagram is changed, and thus the number of internal variables is increased or reduced. Considering an update where the order of the diagram is increased, the new m internal variables are sampled from a distribution $\rho(\mathbf{x})$ and the acceptance ratio R_{ij} is computed as follows:

$$R_{ij} = \frac{p_A}{p_B} \frac{D_{n_i = n_j + m}(i)}{\rho(x_1, ..., x_n) D_{n_j}(j)}, \tag{3.57}$$

where p_A and p_B are the so-called **context factors** that take into account the ways it is possible to transition from state A to state B and thus depends on how the processes are organized (it does not usually reduce to the simple relation $p_A/p_B = 1$) [28]. The opposite process which reduces the order of the diagram from n to $n - m$ is found to be

$$R_{ji} = \frac{1}{R_{ij}}, \tag{3.58}$$

in agreement with the detailed balance requirement.

Employing these rules and carefully crafting prior distributions and acceptance ratios, it is possible to sample a distribution that is equal up to a normalization constant a Green's function in imaginary time formalism, provided that it can be expanded in a similar way to 2.5.

Input : initial configuration $X^{(0)} = \left(\{y\}; \xi_n, x_1, \ldots, x_n\right)$
        update procedures $\{U_1, \ldots, U_k\}$
Output: histogram of $Q(\{y\})$

histogram[];
start from initial diagram $\mathcal{D} \leftarrow X^{(0)}$;
while *not enough samples* **do**
    choose an update U_i from $\{U_1, \ldots, U_k\}$;
    propose a new configuration $X' = \left(\{y'\}; \xi_{n'}, x_{1'}, \ldots, x_{n'}\right)$ according to U_i;
    calculate acceptance ratio R;
    if $R \geq 1$ **then**
        $\mathcal{D} \leftarrow X'$;
    else
        draw random uniform number r;
        if $r \leq R$ **then**
           $\mathcal{D} \leftarrow X'$;
    end
    histogram[$\{y\}$].add(1);
end
return histogram;

Figure 3.3: Diagrammatic Monte Carlo algorithm procedure [5].

The specific updates, however, greatly vary between different systems considering issues such as ergodicity requirements, context factors and the sign problem.

4   Diag MC for the Froehlich Polaron

In chapter 2 we have seen that is possible to perturbatively expand the Matsubara Green's function for the Froehlich polaron in a series of integrals 2.5, while in chapter 3 we have shown that series of this type can be computed using a special type of Markov chain Monte Carlo called Diagrammatic Monte Carlo which is able to take into account all the possible perturbative expansions of the full function and sample from it.

We now focus on the specific Diagrammatic Monte Carlo technique employed for the Froehlich Polaron case. Let us consider the Froehlich Hamiltonian H^{cFr} in the specific case of a single anisotropic electron band and multiple phonon modes. We take into account the non-interacting term

$$H_0^{cFr} = \sum_{\mathbf{k}} \left(\frac{k^2}{2m(\hat{k})} - \mu \right) c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + \sum_{\mathbf{q}j} \omega_{jLO} a_{\mathbf{q}j}^\dagger a_{\mathbf{q}j}, \tag{4.1}$$

where we are only taking into account the electron polaron (the only difference in the hole polaron case would be negative electron energy dispersion). The chemical potential μ is put inside the electron term as a renormalization constant (**fictitious potential renormalization** [24])[28] which can be tuned in order to obtain a non-divergent distribution that decays as an exponential for long τ in the stationary limit.

The interaction term reads:

$$V^{cFR} = \sum_{\mathbf{k},\mathbf{q}j} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} \left(V_{\mathbf{q}j}^{cFr} a_q + V_{\mathbf{q}j}^{*cFr} a_{-\mathbf{q}}^\dagger \right), \tag{4.2}$$

where the interaction vertex V^{cFr} has formula

$$V_{\mathbf{q}j}^{cFr} = \frac{i}{q} \frac{4\pi}{\Omega_0} \left(\frac{1}{2\omega_{jLO} V_{BvK}} \right)^{1/2} \frac{p_j}{\epsilon^\infty}. \tag{4.3}$$

We need a technique to translate the Green's functions defined in chapter 2 to the formalism of Diagrammatic Monte Carlo. To this aim we identify the stationary distribution $Q(\{y\})$ with the two Green's function of interest: the one-electron Green's function $G(\mathbf{k}, \tau)$, where no external phonon lines are present:

$$G(\mathbf{k}, \tau) = \sum_{n=0,2,4,\dots}^{+\infty} \sum_{\xi_n} \int d\tau_1 \cdots \int d\tau_n \int d\mathbf{q}_1 \frac{V_{q_1 j_1}}{(2\pi)^3} \cdots \int d\mathbf{q}_{n/2} \frac{V_{q_{n/2} j_{n/2}}}{(2\pi)^3}$$

$$\times D_n^{\xi_n}(\mathbf{k}, \tau; \tau_1, \dots, \tau_n, \mathbf{q}_1, \dots, \mathbf{q}_{n/2}, j_1, \dots, j_{n/2}), \tag{4.4}$$

and $P(\mathbf{k}, \tau)$, a function that is the sum of the single electron Green's function and all the possible one-electron N-phonons Green's function configurations (weighted accordingly):

$$P(\mathbf{k}, \tau) = G(\mathbf{k}, \tau) + \sum_{N=1}^{+\infty} G^{(N)}(\mathbf{k}\tilde{\mathbf{q}}_1, \dots, \tilde{\mathbf{q}}_N, j_1, \dots, j_N, \tau), \tag{4.5}$$

where $G^{(N)}(\mathbf{k}, \tau)$ is diagrammatically expanded in

$$
\begin{aligned}
G^{(N)}(\cdots) = \sum_{\xi_n} \int d\tau_1 \cdots \int d\tau_n \int d\mathbf{q}_1 \frac{V_{q_1 j_1}}{(2\pi)^3} \cdots \int d\mathbf{q}_{n/2} \frac{V_{q_{n/2} j_{n/2}}}{(2\pi)^3} \\
\times D_n^{\xi_n}(\mathbf{k}, \tilde{\mathbf{q}}_1, ..., \tilde{\mathbf{q}}_N, \tilde{j}_1, ..., \tilde{j}_N, \tau; \tau_1, ..., \tau_n, \mathbf{q}_1, ..., \mathbf{q}_{n/2}, j_1, ..., j_{n/2}).
\end{aligned}
\tag{4.6}
$$

It is now important to distinguish between external variables y and the internal ones $\{x_1, ..., x_n\}$. We identify them as

$$
\begin{aligned}
\{y\} &\to \{\mathbf{k}, \tilde{\mathbf{q}}_1, \tilde{j}_1, ..., \tilde{\mathbf{q}}_N, \tilde{j}_N, \tau\}, \\
n &\to \{0, 2, 4, ...\}, \\
\{x_1, ..., x_n\} &\to \{\tau_1, ..., \tau_n, \mathbf{q}_1, j_1, ..., \mathbf{q}_{n/2}, j_{n/2}\}.
\end{aligned}
\tag{4.7}
$$

The algorithm that computes our simulation must satisfy the ergodicity requirement of Markov chain Monte Carlo: for this reason it is necessary to implement updates that change the external variables $\{y\}$, the internal ones $\{x_i\}$, and the order of the diagram n.

We start by saying that no updates have been implemented in order to sample from the free electron momentum $\mathbf{k}$, which will be fixed for each individual simulation: this was made in order to have a more precise control about the region of $k - space$ that we want to simulate, in particular the $\mathbf{k} = 0$ point where the electron is at the bottom (top) of the conduction (valence) band.

The chemical potential μ is also implemented in the simulations as a normalizing factor, in this way a free electron propagator becomes:

$$
G_0(\mathbf{k}, \tau) \to G_0(\mathbf{k}, \mu, \tau) = e^{-(\epsilon(\mathbf{k})-\mu)\tau},
\tag{4.8}
$$

the effect of which can be easily reversed by rescaling the result.

Due to the finite memory of computers, it is then necessary to fix for each simulation a maximum imaginary time value τ_{max}, a maximum internal order n_{max} (the number of internal phonon propagators would be $n_{max}/2$) and a maximum number of external phonon propagators N_{max}.

Having decided this, it is then necessary to state the electron effective mass on the three main axes m_x^*, m_y^* and m_z^* together with the number of phonon modes that couple to the electron band n_{ph} together with their energy ω_{jLO} and their phonon mode polarities p_j or another equivalent quantity from which their value can be recovered [34].

Other quantities that needs to be specified are the size of the unit cell Ω_0, the size of the Born-von Karman cell V_{BvK}, the optical dielectric tensor (a scalar in the case of cubic materials) ϵ_∞, the number of thermalization steps N_{relax} set to achieve the stationary distribution condition and the number of simulation steps N_{steps}.

It is now important to define how the diagram was effectively modelled in the computer: to this aim three data structures were defined. The **Vertex**: this data structure encapsulates all

the data which describes the vertices of the electron-phonon interaction (also the diagram beginning and end) together with all the relevant information about the associated phonon propagator, its components are:

- τ_i, the imaginary time value of the vertex, 0 and τ for the extrema

- **type**, integer parameter which differentiates between extrema (0), internal phonon vertices (+1 for outgoing phonon propagators and -1 for incoming ones), external phonon vertices (+2 for outgoing phonon propagators and -2 for incoming ones).

- **linked**, integer value which shows to which other vertex the current vertex is linked to and thus the length of the phonon propagator (whether we have an internal or external one).

- $\mathbf{q}_i$, the three momentum components of the free phonon propagator which is created or annihilated at the vertex.

- **index**, the phonon mode which is involved in the electron-phonon interaction at the vertex.

The **Propagator**: this data structure represents the free electron propagators, it is defined by the three electron momentum components $\mathbf{k}_i$.

The **Effective_Mass**: this data structure represents the effective mass $m_i^*(\hat{k})$ of the electron for each electron propagator $\mathbf{k}_i$, it is computed using the formula in 1.78.

At the start of each simulation, an array of dimension $n_{max} + 2N_{max} + 2$ of Vertex data type is created, while an array of dimension $n_{max} + 2N_{max} + 1$ is created for the Propagator and Effective_Mass data type. At the beginning of the simulation all the single elements of the three data types are initialized to default values with the exception of the first 2 Vertex elements and the first Propagator and Effective_Mass element: the latter two are initialized using the $\mathbf{k}$ value given as input and by computing the related effective mass (default to 1 in the case that $\mathbf{k} = 0$), while the τ values of the first two Vertex elements are initialized to 0 and τ respectively, where τ is retrieved using the **Diagram Length** update, which will be explained in detail in the following section.

4.1   Updates

Updates are necessary to obtain a Markov chain for the system which is ergodic, there is no single way to implement them in order to achieve this aim and different choices have been proposed about the way new parameters are proposed and how updates are implemented [24][26], here it is thus shown just a proposal that is sufficient but is nevertheless not necessarily the best or fastest choice.

The first update that is described is the **Diagram Length** update (class I), which updates the length of the last electron propagator: it thus changes the + value of the external variable τ.

Since the last electron propagator is a free electron propagator, which decays exponentially with τ as

$$G_0(\mathbf{k}, \tilde{\mathbf{q}}_1, ..., \tilde{\mathbf{q}}_N, \tau) = \exp\left\{-\left(\frac{p^2}{2m^*(\hat{p})} - \sum_j \omega_{jLO} n_{jph} - \mu\right)(\tau - \tau_{last})\right\}, \quad (4.9)$$

with τ_{last} time of last vertex before the diagram end, and where we have:

$$\mathbf{p} = \mathbf{k} - \sum_{i=0}^{N} \tilde{\mathbf{q}}_i,$$
$$\sum_j n_{jph} = N, \quad (4.10)$$

with N total number of external phonons in the current diagram and j phonon mode index. The weight ratio between two diagrams with length τ and τ' thus becomes:

$$\frac{D_n(\cdots, \tau')}{D_n(\cdots, \tau)} = \frac{e^{-\Delta E(\tau' - \tau_{last})}}{e^{-\Delta E(\tau - \tau last)}}, \quad (4.11)$$

where ΔE is the computed energy in 4.9.
If we thus take an exponential distribution as the proposal distribution for new values of the imaginary time value τ' built as follows

$$\tau' = \tau_{last} - \frac{\log(1 - r)}{\Delta E}, \quad (4.12)$$

with r uniform random variable in $[0, 1]$ we obtain a new value that is always accepted apart from the case where $\tau' > \tau_{max}$. In fact the acceptance ratio is

$$R_{\tau'\tau} = \frac{\exp(\tau; \tau_{last}, \Delta E)e^{-\Delta E(\tau' - \tau_{last})}}{\exp(\tau'; \tau_{last}, \Delta E)e^{-\Delta E(\tau - \tau_{last})}} = \frac{\Delta E e^{-\Delta E(\tau - \tau_{last})}e^{-\Delta E(\tau' - \tau_{last})}}{\Delta E e^{-\Delta E(\tau' - \tau_{last})}e^{-\Delta E(\tau - \tau_{last})}} = 1, \quad (4.13)$$

which is always accepted.
The next update to be described is the **add internal phonon**, a class II update. The new (internal) variables that have to be proposed are the time value of the outgoing vertex τ', the time value of the incoming vertex τ'', the phonon mode index j and the phonon propagator momentum $\mathbf{q}$. If proposing the diagram means that $n + 2 > n_{max}$ maximum allowed internal order, then the update is rejected.
An electron propagator (from the first one to the last for a total of $n + 1$ possible choices) is chosen at random, then the value of τ' is generated using a uniform distribution between τ_{left} and τ_{right} vertices of the chosen free electron propagator.
The phonon mode j is then chosen at random between the ones given in input, and the value

of the second vertex of the phonon propagator τ'' is generated using the following exponential distribution

$$P(\tau''|\tau', j) = \tau' - \frac{\log(1-r)}{\omega_{jLO}}, \tag{4.14}$$

where τ'' is not restrained to any particular free electron propagator. The update is straight-up rejected if $\tau'' > \tau$ length of the full diagram or if it is too close to another vertex ($|\tau'' - \tau_i| < 10^{-9}$).

The phonon momentum values $\mathbf{q}$ are then proposed using a probability $P(\mathbf{q}|\tau', \tau'', j)$ dependent on a gaussian distribution with mean 0 and variance $(\tau'' - \tau')^{-1}$.

The ratio of the two diagrams are

$$\frac{D_{(n+2)}(\mathbf{k}, \tau, ..., \tau', \tau'', \mathbf{q}, ...)}{D_n(\mathbf{k}, \tau, ...)} =$$

$$= \exp\left\{-\left(\sum_i \epsilon(\mathbf{k}_i - \mathbf{q}) - \epsilon(\mathbf{k}_i)\Delta\tau_i - \omega_{LO}(\tau'' - \tau')\right)\right\}|V_{\mathbf{q}j}|^2 d\tau' d\tau'' \frac{d\mathbf{q}}{(2\pi)^3}, \tag{4.15}$$

where the sum over i is extended over all the phonon propagators between the two phonon vertices τ' and τ'', while $\Delta\tau_i$ is computed as $\tau_i - \tau_{i-1}$ where the two extrema are τ' and τ''. The infinitesimal are required because they are not cancelled in the ratio since they are new proposed variables.

The full distribution from which the new values are sampled is written as

$$P(\tau', \tau'', \mathbf{q}) = P(\tau')P(\tau''|\tau', j)P(\mathbf{q}|\tau', \tau'', j) =$$

$$= \underbrace{\frac{1}{\tau_{right} - \tau_{left}}}_{P(\tau')} \underbrace{\omega_{jLO} e^{-\omega_{jLO}(\tau'' - \tau')}}_{P(\tau''|\tau', j)} \underbrace{\left(\frac{\tau'' - \tau'}{2\pi}\right)^{3/2} e^{-\frac{q^2}{2}(\tau'' - \tau')}}_{P(\mathbf{q}|\tau', \tau'', j)}. \tag{4.16}$$

The acceptance ratio thus becomes:

$$R_{add} = \frac{p_A}{p_B} \frac{D_{n+2}(\mathbf{k}, \tau, ..., \tau', \tau'', \mathbf{q}, ...)}{D_n(\mathbf{k}, \tau, ...)P(\tau', \tau'', \mathbf{q})}, \tag{4.17}$$

the context factors p_A and p_B take into account the number of free electron propagators from which a vertex can be generated and the total number of internal phonon propagators which can be removed, their ratio is

$$\frac{p_A}{p_B} = \frac{p_{(rem\,int)}}{p_{(add\,int)}} \frac{n + 2N + 1}{n/2 + 1}, \tag{4.18}$$

for example at order $n = 2$ and 1 external phonon we can choose from $2 + 2 \cdot 1 + 1 = 5$ electron propagators, while to go back to order 2 from order $n + 2 = 4$ we can choose between $2/2 + 1 = 2$ phonon propagators. The two variables $p_{(rem\,int)}$ and $p_{(add\,int)}$ measure

the probability of choosing the add internal and remove internal updates in the main Monte Carlo simulation.

The **remove internal phonon** is also a class II update, no new parameters are proposed but instead a random phonon propagator with vertices at τ' and τ'' and momentum $\mathbf{q}$ is chosen to be removed. The update is automatically rejected if the internal order $n = 0$ and is accepted with acceptance ratio $R_{rem} = 1/R_{add}$ calculated for the add internal phonon update.

4.2   Collected quantities and exact estimators

5 Acknowledgement

Appendix

References

[1] Craig Pryor. Predicted band structures of iii-v semiconductors in the wurtzite phase. *Phys. Rev. B*, 81, 08 2009.

[2] Jozef Devreese and Alexandre Alexandrov. Froehlich polaron and bipolaron: Recent developments. *Reports on Progress in Physics*, 72:066501, 05 2009.

[3] Jun Da Ng and Aaron Danner. Pseudopotential form factors and electronic band structures for alas, alp, bas, bp, 3c-sic, and cubic-gan. *Physica Scripta*, 96(5):055801, 2021.

[4] James Gubernatis, Naoki Kawashima, and Philipp Werner. *Quantum Monte Carlo Methods*. Cambridge University Press, 2016.

[5] Thomas Hahn. Diagrammatic quantum monte carlo for the fröhlich polaron, 2017.

[6] Felix Bloch. Quantum mechanics of electrons in crystal lattices. *Z. Phys*, 52:555–600, 1928.

[7] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.

[8] Jos Thijssen. *Computational physics*. Cambridge university press, 2007.

[9] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[10] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 11 1965.

[11] GP Kerker. Non-singular atomic pseudopotentials for solid state applications. *Journal of Physics C: Solid State Physics*, 13(9):L189, 1980.

[12] Paul Ziesche, Stefan Kurth, and John P Perdew. Density functionals from lda to gga. *Computational materials science*, 11(2):122–127, 1998.

[13] Herbert Fröhlich. Electrons in lattice fields. *Advances in Physics*, 3(11):325–361, 1954.

[14] Hermann Haken. *Quantum Field Theory of Solids. An Introduction*. 1976.

[15] Alexandre S Alexandrov and Jozef T Devreese. *Advances in polaron physics*, volume 159. Springer, 2010.

[16] Gerald D Mahan. *Many-particle physics*. Springer Science & Business Media, 2013.

[17] Satoru J Miyake. The ground state of the optical polaron in the strong-coupling case. *Journal of the Physical Society of Japan*, 41(3):747–752, 1976.

[18] Anna Miglio, Véronique Brousseau-Couture, Emile Godbout, Gabriel Antonius, Yang-Hao Chan, Steven G Louie, Michel Côté, Matteo Giantomassi, and Xavier Gonze. Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap. *npj Computational Materials*, 6(1):167, 2020.

[19] Bogdan Guster, Pedro Melo, Bradley AA Martin, Véronique Brousseau-Couture, Joao C de Abreu, Anna Miglio, Matteo Giantomassi, Michel Côté, Jarvist M Frost, Matthieu J Verstraete, et al. Fröhlich polaron effective mass and localization length in cubic materials: Degenerate and anisotropic electronic bands. *Physical Review B*, 104(23):235123, 2021.

[20] Xavier Gonze and Changyol Lee. Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Physical Review B*, 55(16):10355, 1997.

[21] Henrik Bruus and Karsten Flensberg. *Many-body quantum theory in condensed matter physics: an introduction*. Oxford university press, 2004.

[22] Wolfgang Nolting and William D Brewer. *Fundamentals of many-body physics*, volume 3. Springer, 2009.

[23] Murray Gell-Mann and Francis Low. Bound states in quantum field theory. *Phys. Rev.*, 84:350–354, 10 1951.

[24] AS Mishchenko, NV Prokof'Ev, A Sakamoto, and BV Svistunov. Diagrammatic quantum monte carlo study of the fröhlich polaron. *Physical Review B*, 62(10):6317, 2000.

[25] Andrei S Mishchenko. Diagrammatic monte carlo method as applied to the polaron problems. *Physics-Uspekhi*, 48(9):887, 2005.

[26] Thomas Hahn, Sergei Klimin, Jacques Tempere, Jozef T Devreese, and Cesare Franchini. Diagrammatic monte carlo study of fröhlich polaron dispersion in two and three dimensions. *Physical Review B*, 97(13):134305, 2018.

[27] Nicholas Metropolis and Stanislaw Ulam. The monte carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949.

[28] Holger Fehske, Ralf Schneider, and Alexander Weisse. *Computational many-particle physics*, volume 739. Springer, 2007.

[29] Federico Becca and Sandro Sorella. *Quantum Monte Carlo approaches for correlated systems*. Cambridge University Press, 2017.

[30] Makoto Matsumoto and Takuji Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 8(1):3–30, 1998.

[31] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[32] Emanuel Gull, Andrew J Millis, Alexander I Lichtenstein, Alexey N Rubtsov, Matthias Troyer, and Philipp Werner. Continuous-time monte carlo methods for quantum impurity models. *Reviews of Modern Physics*, 83(2):349–404, 2011.

[33] Kris Van Houcke, Evgeny Kozik, Nikolay Prokof'ev, and Boris Svistunov. Diagrammatic monte carlo. *Physics Procedia*, 6:95–105, 2010.

[34] Pedro Miguel MC de Melo, Joao C de Abreu, Bogdan Guster, Matteo Giantomassi, Zeila Zanolli, Xavier Gonze, and Matthieu J Verstraete. High-throughput analysis of fröhlich-type polaron models. *npj Computational Materials*, 9(1):147, 2023.