Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Physics

Quantum Monte Carlo Study of Effective Masses in a Polarized Unitary Fermi Gas

Supervisor:
Prof. Pierbiagio Pieri

Co-supervisor:
Dr. Gianluca Bertaina

Submitted by:
Nitya Cuzzuol

Academic Year 2021/2022

ABSTRACT

Ultracold gases provide an ideal platform for quantum simulations of many-body systems, due to the high controllability of various parameters ranging, for instance, from their interaction to the population of particles. Quite generally, the interaction in ultracold gases experiments has a short range compared to the average interparticle distance and can thus be described in terms of a single physical parameter, namely the (s-wave) scattering length. A regime of the interaction which has been the focus of most experimental and theoretical works on ultracold fermionic gases is the so-called unitary regime, corresponding to a diverging scattering length. In such a regime, interactions are sufficiently strong to invalidate perturbative methods, thus requiring theoretical approaches, such as Quantum Monte Carlo simulations, that are intrinsically non-perturbative. In this thesis, we study with Quantum Monte Carlo simulations a two-component gas of fermionic atoms at zero temperature in the unitary regime. When the populations of the two spin components are equal, the ground state of the unitary Fermi gas is known to be a superfluid of Cooper pairs. For large imbalances between the two spin populations, i.e., for strongly polarized systems, the ground state is instead a normal Fermi liquid. In this work, we are interested in studying how the effective masses for the two kinds (spin up and spin down) of quasi-particles of the Fermi liquid evolve as the polarization is progressively reduced from full to lower values. A recent theoretical work, based on alternative diagrammatic methods, has indeed suggested that such effective masses should diverge at a critical polarization. To independently verify such predictions, we perform Variation Monte Carlo (VMC) calculations of the energy based on Jastrow-Slater wavefunctions after adding or subtracting a particle with a given momentum to a full Fermi sphere. In this way, we determine the quasi-particle dispersions, from which we extract the effective masses for different polarizations. The resulting effective masses turn out however to be quite close to the non-interacting values, even though some evidence of an increase for the effective mass of the minority component appears close to the predicted value for the critical polarization. Furthermore, preliminary results obtained for the majority component with the Fixed-node Diffusion Monte Carlo (DMC) method seem to indicate that DMC could lead to an increase of the effective masses in comparison with the VMC results. Finally, we point out further improvements of the trial wave-function and boundary conditions that would be necessary in future simulations to draw definite conclusions on the effective masses of the polarized unitary Fermi gas.

CONTENTS

CHAPTER 1

INTRODUCTION

The foundation of the work in this thesis is the application of the Quantum Monte Carlo method to simulate a quantum system that can be described as a Fermi liquid.
The system under consideration is a two-component gas of interacting fermionic atoms at zero temperature in a specific condition of their interaction which is called *unitarity regime*. The two components of the gas correspond to two different isospin states of the atoms, which can be treated effectively as the two independent states of a spin-1/2 system. In the unitarity regime the interaction between the two components (which is of short range under standard experimental conditions) is tuned in such a way that the s-wave scattering length is divergent. In addition, the two spin populations can be controlled independently. When the populations of the two spin components are equal, the ground state of the gas is known to be a superfluid of Cooper pairs. For large imbalances between the two spin populations, i.e., for strongly polarized systems, the ground state of the unitary Fermi gas is instead a normal Fermi liquid.
A recent theoretical work [1] has pointed out that as the polarization is progressively reduced, the effective masses of spin-up and spin-down quasi-particle should diverge at a critical polarization. This critical polarization corresponds to a quantum critical point at which the polarized normal phase gives way to an unusual superfluid phase with finite polarization, with Cooper pairs condensed in a state with finite center of mass momentum. This kind of superfluid phase, which was first proposed many years ago independently by Fulde and Ferrell [2] and Larkin and Ovchinnikov [3], has been searched for a long time in condensed matter systems (with robust evidence of this phase coming only recently [4][5][6][7][8][9][10]), as well as, more recently, with ultracold gases (where robust evidence of this phase is still missing). The prediction of very large effective masses in the proximity of the transition to such a phase is thus of outmost interest. This motivates the work of the present thesis.
Specifically, to independently verify such recent predictions, we will perform Variational Monte Carlo (VMC) calculations of the energy based on Jastrow-Slater wavefunctions. This wavefunction, as it will be described later on, contains two variational parameters that are optimized in our simulations. This way, we will determine the quasi-particle dispersions, from which the effective masses can be extracted for different polarizations. The resulting effective masses will turn out to be quite close to the non-interacting values, even though some evidence of an increase of the effective mass of the minority component close to the predicted value for the critical polarization will be found. Furthermore, we will present preliminary results obtained for the majority component with the Fixed-node Diffusion Monte Carlo (DMC) method which seem to indicate that DMC could lead to an increase of the effective masses in comparison with

the VMC results.
Finally, we will point out some improvements of the trial wave-function and boundary conditions that will be necessary in future simulations to draw definite conclusions on the effective masses of the polarized unitary Fermi gas.

The thesis is organized as follows.
After the present introduction, Chapter 2 describes the theory of Fermi liquids, which, despite having a phenomenological origin, has a strong experimental support, as well as a microscopic foundation. This theory provides the framework for interpreting our simulations. Chapter 3 gives an introduction to ultracold Fermi gases, some aspects of the quantum theory of scattering relevant to the present work, as well as a more in-depth discussion of polarized Fermi systems. Chapter 4 describes Monte Carlo methods, focusing in particular on the Variational Monte Carlo and Fixed-node Diffusion Monte Carlo methods used in our simulations. Chapter 5 provides more details on their specific implementation for the system of interest to us. The main results are presented and discussed in Chapter 6, while Chapter 7 reports our conclusions. The Appendices contain more technical material on Fermi liquid theory (appendices A and B) and Monte Carlo methods (appendices C and D).

CHAPTER 2

FERMI LIQUID THEORY

This chapter will go over the theory of the physical system that underpins our simulations, namely the theory of Fermi liquids. The concept of quasiparticle will be introduced first, followed by a brief review of the ideal Fermi gas results. We will continue by including adiabatically interactions and by discussing what happens to the system. Then we will go over a microscopic description of this theory and a generalization for imbalanced spin populations. In conclusion we will discuss the theory breakdown.

2.1 Elementary excitations and quasiparticles

In general, the quantum properties of a system determine its macroscopic observations only at very low temperature. When the de Broglie wavelength becomes of the order of interatomic distances, quantum effects act in. Even though classical physics predicts that all liquids will solidify before reaching this point, some systems, like liquid helium, can remain liquid in presence of specific interactions.
When a quantum system is isolated, the state is characterized by its energy levels, corresponding to stationary states. Otherwise if the system is connected to a background, energy and particles can be exchanged and the subsystem can now move in any energy state. Even if in particular simple cases one could completely determine the system by calculating the partition function and related thermodynamic quantities, in general this is not possible. The infinity of available states leads to non-calculable functions. So one has to abandon this strategy, except for the case of weak interactions, which allows for perturbation expansions. However, when the temperature is near absolute zero, there is only a little amount of available excited states. In such a situation one can describe quite generally the system, regardless of the magnitude and specific features of interaction.
Now one needs a way to describe such excited states of a macroscopic sample. In quantum mechanics every weakly excited states can be considered as a set of elementary excitations, that act as quasiparticles in motion inside the volume of the system with some specific energies and momenta. As an example, one can consider the lattice vibrations, where one can exploit the quantum-mechanical correspondence principle [11] to introduce a model of oscillators, and in this way one can describe the spectrum of excitations. This concept states that an oscillator of frequency $\omega(\mathbf{p})$ and momentum $\mathbf{p}$, or the equivalent plane wave, can be connected to a set of moving particles with energies and momenta determined by wave properties. If the system

is ideal, these new particles are free, and the energy levels are that of an ideal gas.
Then, we will assume that any energy level can be obtained as the sum of energies of a certain type of quasi-particles or elementary-excitations, moving with momentum $\mathbf{p}$ and energy $\varepsilon(\mathbf{p})$, in the volume occupied by the system. In general the dispersion of these particles is different from the expression for its original components. This new type of excitations results from interaction of the system as a whole and does not correspond to its separate constituents. They can be bosonic or fermionic, depending on their spin (integer or half-integer respectively).
Transitions between states with different numbers of excitations are possible when non-ideal terms are considered, and interactions between these excitations become relevant, resulting in processes of scattering and creation of new quasi-particles. When the amplitude of these non-ideal effects increases, i.e., when the temperature increases, so does the number of excitations and the importance of their interactions. Transitions between states by the new particles are possible because they represent the superposition of a large number of exact stationary states with a finite energy spread. This causes the packet to spread more, resulting in attenuation of the excitation. This type of description is only possible if the packet's width, which determines its attenuation, is small in comparison to the packet's energy. This process can be thought of as interactions between quasi-particles that satisfy energy and momentum conservation. Since they can be identified as decay and scattering processes, they are important if the energy and quasi-particle number are sufficiently large. When the system's energy is low, i.e. when the temperature is low, both effects are negligible, and we can describe the system as an ideal gas of quasi-particles [12].

2.2 Fermi gas

Even though interactions of various types exist in a normal metal, the ideal Fermi gas model can describe some qualitative features of the system, such as the presence of a well-defined Fermi surface. It is then useful to briefly introduce the results for a non-interacting translational invariant system.
The single-particle eigenstates are plane waves

$$\langle \mathbf{r}|\mathbf{k}\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.1}$$

where V is the volume of the system, the corresponding energy is

$$\epsilon(\mathbf{k}) = \epsilon_{\mathbf{k}} = \frac{\hbar^2\mathbf{k}^2}{2m} \tag{2.2}$$

(in general we will set units of $\hbar = 1$).
The ground state of an N-particle system at zero temperature is the well-known Fermi sea: in momentum space fermions occupy all states in increasing momentum order, up to a critical value Fermi wavevector

$$k_F = \left(\frac{6\pi^2 N}{V}\right)^{1/3} \tag{2.3}$$

where N/V is the density of particles per spin component. In thermodynamic limit these states are contained in a sphere of radius k_F, the Fermi sphere. All the other states are empty. The energy of the last occupied state is the so-called Fermi energy and it corresponds to the chemical potential in the limit of $T = 0$.

Every energy eigenstate of the non-interacting system can be described in terms of the occupancy number of the levels, indexed by spin and momentum. In the ground state they are all one up to the Fermi momentum. So the state can be written as

$$|\Psi\rangle = |n_{\mathbf{p}_1\sigma_1}, n_{\mathbf{p}_2\sigma_2}, \dots\rangle \ . \tag{2.4}$$

For the ground state, $n_{\mathbf{p}\sigma} = 1$ if $p < k_F$ and zero otherwise.
When excitations are introduced, one can consider adding a particle with wavevector $k > k_F$ ($\delta n_{\mathbf{k}} = 1$), so energy greater then the chemical potential, or destroying particle with wavevector $k < k_F$ ($\delta n_{\mathbf{k}} = -1$), so energy smaller then the chemical potential. If one wishes to keep the number of particles constant, excitations correspond to the extraction of fermions from inside the Fermi sphere. Then one fermion with momentum $k > k_F$ and a vacancy (or hole) with momentum $k < k_F$ are generated simultaneously.
The process of excitations can be more conveniently described using particle and hole creation operators (or particle destruction) acting on N-fermions ground state by means of second quantization formalism

$$\begin{aligned} |\mathbf{k}; N+1\rangle &= a^\dagger_{\mathbf{k}}|\Psi_0(N)\rangle \\ |\mathbf{k}; N-1\rangle &= a_{-\mathbf{k}}|\Psi_0(N)\rangle \end{aligned} \tag{2.5}$$

for $k > k_F$ (particles) or $k < k_F$ (holes).
However we are interested in a more complicated situation, and including interactions between particles is necessary. Then in next section we will discuss the Landau's theory of Fermi liquids.

2.3 Fermi liquid

Let us now follow [11] and [12] in considering a system of interacting fermionic particles, and inspect the theory of Fermi liquids constructed by Landau within its works [13] [14] [15]. This theory attempts to construct its analysis of the interacting system by leveraging similarities with the ideal case. These two models are connected by switching on interactions infinitely slowly, assuming that the ground state remains stable throughout this process. This is the assumption of adiabaticity, property that, together with the Pauli exclusion principle, is at the basis of the theory. The possibility to describe interacting particles by means of similar non-interacting one relies on the stability of the weak excitations of the system, they have a lifetime so long to be approximate as stable, because of the small phase space available for decay. Landau named this long lived excitations quasiparticles. As a result, each quantum state of the interacting system is in one-to-one correspondence with the non-interacting system.
^{3}He at low temperature is the perfect example for realization of this theory, ^{3}He is a rare spin-1/2 isotope formed by one neutron, two protons and two electrons. The characteristics that make it particularly suitable are its isotropy, the absence of crystalline order, and its neutrality of charge, so that the coulomb interaction is not present.

2.3.1 Elementary excitations

The notion of elementary excitation makes sense only close to the Fermi moment, in fact only there elementary excitations have a low attenuation compared to their energy. This depends on the fact that the attenuation is the result of decay processes and collisions which have a probability proportional to the distance from the Fermi surface. In particular, if the excitation energy is high compared to the liquid temperature, decay is the first cause of attenuation.

Its magnitude is proportional to the probability of decay, which is of the order of $(k - k_F)^2$. Furthermore, in the case of finite temperature, the probability of decay is proportional to T^2, implying that the theory is applicable to low temperatures. If this is the case, an analogy with the Fermi gas can be drawn. When a set of quasi particles is considered around a Fermi sphere, the number of quasi particles equals the number of original particles, but now the final state has an unknown energy E_0.
If the interactions are turned on infinitely slowly, the ground state of the non-interacting system $|\Psi_0(N)\rangle$ moves into the ground state of the interacting one $|\Psi(N)\rangle$. If we repeat the same procedure but with the addition of one fermion above the Fermi surface, this excited state can become unstable and decay, emitting particle-hole pairs that dissipate energy. To avoid this, the particle's energy must be close to the Fermi energy, which results in a long lifetime, that could become infinite if the excitation is on the Fermi surface. In practice, this limits the validity of the theory to a small energy shell around the Fermi energy.
Adiabatic evolution conserves the momentum of the quasi-particle state, that evolves smoothly into $|\Psi_{\mathbf{p}',\sigma'}\rangle$, where now the occupation number is different from zero for $p < k_F$ and $\mathbf{p} = \mathbf{p}'$, the quasiparticle's momentum. This state has total momentum $p' > k_F$, where the Fermi momentum is unchanged after switching on the interactions, and it satisfies the same relation with density as eq.(2.3).
The energy to create a single quasiparticle is

$$E^{(0)}_{\mathbf{p}'} = (E(\mathbf{p}') - E_0) \tag{2.6}$$

where we indicated one single quasi particle in the absence of other excitations with a superscript (0). We can express this energy per particle in grand canonical ensemble subtracting the chemical potential

$$\varepsilon^{(0)}_{\mathbf{p}'} = E^{(0)}_{\mathbf{p}'} - \mu \tag{2.7}$$

Similarly one can describe the quasi-hole state, where a particle inside Fermi sphere is removed, and the energy is

$$\bar{\varepsilon}^{(0)}_{\mathbf{p}'} = -E^{(0)}_{\mathbf{p}'} + \mu = -\varepsilon^{(0)}_{\mathbf{p}'} \tag{2.8}$$

that is still positive since $\varepsilon^{(0)}_{\mathbf{p}'}$ is negative for $p' < k_F$.
The previous description deals only with one single quasiparticle, but when the temperature is finite there is a gas of excitations around the Fermi surface. They interact mutually, but, because they must satisfy the exclusion principle, there is a fixed momentum they can exchange during collisions: on the Fermi surface, this momentum is zero. As a result, these particles are forced to be stationary in their states, and the occupation number is conserved.

2.3.2 Energy functional

When interactions are enabled, the total momentum is the same as for ideal Fermi gas states. Interactions between particles within the Fermi sea and between the excited particle and the Fermi sea, produce Landau quasiparticle effects. The interactions have no effect on the value of k_F, which imposes a lower limit on the allowed momentum of the quasiparticle. Luttinger [16] has shown that the volume enclosed by the Fermi surface in momentum space is not affected by the interaction, but the distribution of particles in k-space changes. Also the total energy is different and it is not the simple sum of quasiparticles' energies. It becomes a functional of the distribution function.

Quasiparticle energy

Landau's theory describes interactions using a self-consistent field, acting on one quasiparticle, generated by all the others. So the energy of the system is a functional of the distribution function of quasi-particles. Then the energy of a single quasi particle is the variational derivative of the total energy with respect to the distribution function. Starting from the equilibrium distribution of quasi particles as $n_0(\mathbf{k})$, and considering changes in quasiparticle occupation number of the the form $n_0(\mathbf{k}) \to n_0(\mathbf{k}) + \delta n(\mathbf{k})$, the corresponding change in energy (in unit volume) is

$$\delta E = \int \varepsilon(\mathbf{p}) \delta n \frac{d\mathbf{p}}{(2\pi)^3} \tag{2.9}$$

the quantity $\varepsilon(\mathbf{p})$ is the variation of the energy of the system when one quasiparticle of momentum $\mathbf{p}$ is added to the system, this is the quantity that defines the behaviour of one quasiparticle in a field of other quasiparticles. It is also a functional of the distribution function, so it depends on the distribution of all particles of the liquid. So in this case the presence of other particles is felt not only through the interaction energy but also the kinetic energy is strongly influenced by the other particles.
In general, one can express $\varepsilon(\mathbf{p})$ with a spin dependence, replacing the previous quantities in eq.(2.9) with matrices and spin operators, that, in an isotropic and homogeneous liquid, where the spin operator is a degeneracy constant, are transformed in a quantity proportional to the unit matrix. The spin degeneracy can always be re-conducted to spin-1/2 quasiparticles, the difference, in case of particles of any spin, is just the number of branches of the function $\varepsilon(\mathbf{p})$ for spin-1/2 quasiparticles [12].
Let us replace now $\varepsilon(\mathbf{p})$ and n with the respective operators $\hat{\varepsilon}(p)$ and $\hat{n}$, and let us introduce the trace operation as $\mathrm{Tr}[\cdot]$ in eq.(2.9)

$$\delta E = \mathrm{Tr} \int \hat{\varepsilon}(\mathbf{p}) \delta \hat{n} \frac{d\mathbf{p}}{(2\pi)^3} = \sum_{\alpha,\,\beta} \int \varepsilon_{\alpha\beta} \delta n_{\beta\alpha} \frac{d\mathbf{p}}{(2\pi)^3} \tag{2.10}$$

We can consider the expansion of the energy in power of $\delta\hat{n}$, and rewrite the previous equation as

$$\delta E = \sum_{\alpha,\,\beta} \int \varepsilon^{(0)}_{\alpha\beta}(\mathbf{p}) \delta n_{\beta\alpha} \frac{d\mathbf{p}}{(2\pi)^3} + \frac{1}{2} \sum_{\alpha,\,\beta} \iint f_{\alpha\beta,\gamma\delta}(\mathbf{p},\mathbf{p}') \delta n_{\beta\alpha}(\mathbf{p}) \delta n_{\delta\gamma}(\mathbf{p}') \frac{d\mathbf{p}'}{(2\pi)^3} \frac{d\mathbf{p}}{(2\pi)^3} \tag{2.11}$$

where the first order coefficient describes the energy of an isolated quasiparticle. The function f is the second order coefficient, which corresponds to the second functional derivative with respect to $\delta\hat{n}$ of the energy per unit volume, and is symmetric under permutation of the pair momenta and spins. At the second order of the expansion, the energy depends on the variation of the quasiparticle occupancies.
If $\hat{\varepsilon}^{(0)}(\mathbf{p})$ describes the equilibrium energy for a single quasi-particle, then, for small deviation from equilibrium, we can compute the variation of total energy and obtain the quasi-particle energy as

$$\hat{\varepsilon}(\mathbf{p}) = \hat{\varepsilon}^{(0)}(\mathbf{p}) + \delta\hat{\varepsilon}(\mathbf{p}) = \hat{\varepsilon}^{(0)}(\mathbf{p}) + \mathrm{Tr} \int \hat{f}(\mathbf{p},\mathbf{p}') \delta\hat{n}(\mathbf{p}') \frac{d\mathbf{p}'}{(2\pi)^3} \tag{2.12}$$

where the first term is the energy of a bare quasi–particle, while the second one represents the contribution from interactions between quasiparticles, that is an energy change induced by polarization of the Fermi sea.

Quasiparticle distribution

When interactions are turned on adiabatically in an ideal system, the same energy levels classification as for the ideal gas applies to the system, so we can calculate the entropy as we would do for a non-interacting gas, and we can use the same dependence on the distribution function

$$S(\hat{n}) = -k_B \mathrm{Tr} \int \left[\hat{n} \ln \hat{n} + (1-\hat{n}) \ln (1-\hat{n})\right] \frac{d\mathbf{p}'}{(2\pi)^3} \tag{2.13}$$

Eqs. (2.11) and (2.13), can be combined to obtain the free energy, from which one can extract the full thermodynamics,

$$\begin{aligned} F(\hat{n}) =& (E(\hat{n}) - \mu N) - TS(\hat{n}) \\ =& (E_0 - \mu N) + \mathrm{Tr} \int \hat{\varepsilon}^{(0)}(\mathbf{p}) \delta\hat{n} \frac{d\mathbf{p}}{(2\pi)^3} + \frac{1}{2} \mathrm{Tr} \iint \hat{f}(\mathbf{p}, \mathbf{p}') \delta\hat{n}(\mathbf{p}) \delta\hat{n}(\mathbf{p}') \frac{d\mathbf{p}'}{(2\pi)^3} \frac{d\mathbf{p}}{(2\pi)^3} \\ &+ T k_B \mathrm{Tr} \int \left[\hat{n} \ln \hat{n} + (1-\hat{n}) \ln (1-\hat{n})\right] \frac{d\mathbf{p}'}{(2\pi)^3} \end{aligned} \tag{2.14}$$

In thermal equilibrium the free energy is stationary with respect to small changes in quasiparticle occupancy, so one can compute the equilibrium momentum distribution by setting $\delta F = 0$

$$\delta F = \mathrm{Tr} \int \delta\hat{n}(\mathbf{p}) \left[\hat{\varepsilon}(\mathbf{p}) + k_B T \ln \left(\frac{\hat{n}(\mathbf{p})}{1-\hat{n}(\mathbf{p})}\right)\right] + O(\delta\hat{n}^2(\mathbf{p})) = 0 \tag{2.15}$$

just neglecting the $\delta\hat{n}(\mathbf{p})$ inside the logarithm at the price of an $O(\delta\hat{n}^2(\mathbf{p}))$ error. The resulting distribution function is

$$\hat{n}(\hat{\varepsilon}) = \frac{1}{e^{(\hat{\varepsilon}-\mu)/k_B T} + 1} \ . \tag{2.16}$$

Since $\hat{\varepsilon}$ is a functional of $\hat{n}$, then we have a quite complicated self-consistent relation 2.16, and a temperature dependence appears in $\hat{\varepsilon}$ through $\hat{n}$.

2.3.3 Effective mass

When energy and temperature are sufficiently low, we can use the distribution function of a Fermi gas and compute the energy functional $\hat{\varepsilon}$. In this case, assuming also spin independence, the distribution function becomes equivalent to a normal Fermi-Dirac distribution. With this assumptions, and restricting to momenta near k_F, we can write the energy as a series expansion of $(p - k_F)$. Taking only the lowest term, it is

$$\varepsilon^{(0)}(\mathbf{p}) - \varepsilon_F \approx v_F (p - k_F) \tag{2.17}$$

where $\varepsilon_F = \mu(T = 0)$. The constant v_F represents the velocity of the excitations on the Fermi surface. It is possible to express this quantity exploiting the analogy with the ideal Fermi gas

$$v_F = \frac{k_F}{m} \rightarrow v_F = \frac{k_F}{m^*} \tag{2.18}$$

Landau's Fermi liquid theory connects the description of the interacting system to a non interacting one just introducing a renormalization of the non-interacting quantities. In this way the effects of interactions on quasiparticles energies is taken into account. We are introducing here an effective mass of the quasi-particles, this quantity is different in general from the bare mass of the particle.

2.3.4 Interaction function

We are interested in finding the connection between the function $\hat{f}$ and the effective mass. In order to find this relation, and many others, we have to analyse the interaction function. If spin is conserved, the interaction is invariant under spin rotation, and it is convenient to express the $\hat{f}$ function in terms of its symmetrical and anti-symmetrical components. One can write a general rotational invariant form for $\hat{f}$ introducing the Pauli matrices $\boldsymbol{\sigma}$ as in [11]

$$f_{\alpha\beta,\gamma\eta}(\mathbf{p},\mathbf{p}') = f^s(\mathbf{p},\mathbf{p}')\delta_{\alpha\beta}\delta_{\gamma\eta} + f^a(\mathbf{p},\mathbf{p}')\sigma_{\alpha\beta}\sigma_{\gamma\eta} \tag{2.19}$$

from which one can take the diagonal components. Now, since $\delta\hat{n}(\mathbf{p}')$ is considerably different from zero only in a narrow range of $\mathbf{p}'$ values near the Fermi wavevector, and the same for $\mathbf{p}$, it is possible to approximate the interaction function as the value assumed when $|\mathbf{p}| = |\mathbf{p}'| = k_F$. Considering an isotropic Fermi liquid, this function depends only on the angle θ between vectors, so we can write

$$f^{s,a}(\mathbf{p},\mathbf{p}') = f^{s,a}\left(\frac{\mathbf{p}\cdot\mathbf{p}'}{k_F^2}\right) = f^{s,a}\left(\cos(\theta)\right) \tag{2.20}$$

that is also convenient to transform into a dimensionless form

$$F^{s,a} = \frac{k_F m^*}{\pi^2} f^{s,a} = N^*(0) f^{s,a} \tag{2.21}$$

where the constant corresponds to the quasiparticle density of states of the interacting liquid $N^*(\varepsilon)$ at the Fermi energy. One can also expand F in Legendre polynomial as

$$F^{s,a}(\mathbf{p},\mathbf{p}') = \sum_{l=0}^{\infty} F_l^{s,a} P_l\left(\cos(\theta)\right) \tag{2.22}$$

These coefficients $F_l^{s,a}$ are the so called Landau parameters, they will be used to renormalize the non-interacting quantities.

Interaction coefficients

Following [17], one can visualize the Landau Fermi liquid as a deformable sphere. The Fermi sphere changes shape when the density or magnetization of the fluid is modified, or if a current flows. These deformations act back on the quasiparticles via the Landau interactions and change the quasiparticle energies. To examine the feedback effects of interactions, let us suppose that an external potential, or field, is applied to induce a polarization of the Fermi surface. The applied field couples to a conserved quantity, which is unchanged by interactions. This means that for any quasiparticle configuration $\delta\hat{n}(\mathbf{p})$, the energy associated with the application of the external field is unaffected by interactions, ensuring that the coupling to the external field is the same as that of non-interacting particles.
We can establish a connection between the size of the energy variation and the feedback effect, as discussed in Appendix A. The application of an external potential (or field) adds a term $\delta\hat{\varepsilon}^{(0)}$ to the quasi-particle energy variation, which is due to the external field

$$\delta\hat{\varepsilon} = \mathrm{Tr}\int \hat{f}(\mathbf{p},\mathbf{p}')\delta\hat{n}(\mathbf{p}')\frac{d\mathbf{p}'}{(2\pi)^3} + \delta\hat{\varepsilon}^{(0)} \ . \tag{2.23}$$

This variation reflects on quasiparticle distribution as

$$\hat{n} = f(\hat{\varepsilon}^{(0)} + \delta\hat{\varepsilon}) \approx f(\hat{\varepsilon}^{(0)}) + f'(\hat{\varepsilon}^{(0)})\delta\hat{\varepsilon} \tag{2.24}$$

where here f indicates the Fermi distribution.
After relaxing the spin dependence and considering the case of a particular multipole symmetry in the quasiparticle potential, one can obtain the following bare change of quasiparticle energy

$$\delta\varepsilon^{(0)}(\mathbf{p}) = v_l Y_{lm}(\hat{\mathbf{p}}) \tag{2.25}$$

where Y_{lm} is a spherical harmonic, while v_l is the magnitude of this variation. Since we are assuming that the system maintains the same symmetry after switching on interactions, the total energy variation $\delta\hat{\varepsilon}$ has the same multipole symmetry of the $\delta\varepsilon^{(0)}$, while the magnitude changes from v_l to t_l,

$$\delta\varepsilon(\mathbf{p}) = t_l Y_{lm}(\hat{\mathbf{p}}) \ . \tag{2.26}$$

This leads to the variation of the occupation number as in eq.(2.24), that, after some calculations, leads to the total energy variation

$$\delta\varepsilon(\mathbf{p}) = v_l Y_{lm}(\hat{\mathbf{p}}) - \frac{t_l}{2l+1} F_l^s Y_{lm}(\hat{\mathbf{p}}) = t_l Y_{lm}(\hat{\mathbf{p}}) \tag{2.27}$$

that permits to extract the following relation

$$t_l = \frac{v_l}{1 + \frac{F_l}{2l+1}} \tag{2.28}$$

Then one can see that depending on the sign of F_l, the resulting magnitude of the energy variation is different. In particular, if $F_l > 0$, a negative feedback occurs which causes the response to be suppressed. Instead, if $F_l < 0$, a positive feedback occurs which causes the response to be enhanced. The critical value $F_l = -(2l+1)$ generates an instability to deformation of the Fermi surface, the so-called Pomeranchuk instability [18].

2.3.5 Effective mass and quasi-particles interactions

One important quantity that is renormalized because of the effects previously described is the mass, that becomes an effective mass of quasiparticle. When the quasiparticles fluid is set in motion at velocity $\mathbf{u}$, this induces a dipolar deformation of the Fermi surface, which is fed-back via the dipolar Landau parameter F_1^s. This process induces a renormalization of the effective mass. As the fermion moves through the medium, the backflow of the surrounding fluid modifies it effective mass according to the relation

$$m^* = m\left(1 + \frac{F_1^s}{3}\right) \ . \tag{2.29}$$

We can ascribe a particle current $v_F = k_F/m^*$ to each quasiparticle, that can be written as

$$v_F = \frac{k_F}{m^*} = \frac{k_F}{m\left(1 + F_1^s/3\right)} = \frac{k_F}{m} - \frac{k_F}{m}\left(\frac{F_1^s/3}{1 + F_1^s/3}\right) \ . \tag{2.30}$$

The first term is the bare current associated to a bare particle, while the second term is the backflow of the surrounding Fermi sea. It is a feedback response to the dipolar distortion of the Fermi surface which develops in the presence of a current. An increase of the effective mass has the effect of compressing the spacing between energy levels, and this increases the density of states and then also the number of excited quasiparticles at a given temperature [17].
In order to prove the relation 2.29 between bare mass and effective mass, one can follow the

original method by Landau [11], exploiting the Galilean invariance, and set the equivalence of momentum for unit volume of liquid and density of its mass flux, as one can see in Appendix B. Otherwise one can use the previous general description of the feedback effect [17] and introduce a vector potential $\mathbf{A}$, which is conjugate to the momentum (a conserved quantity). Then one has to apply the relation in eq.(2.28) to the energy variation

$$\delta\varepsilon^{(0)}(\mathbf{p}) = -\frac{\mathbf{p}}{m}\cdot\mathbf{A}\ . \tag{2.31}$$

Note that by using eq.(2.21), eq.(2.29) can also be written as

$$m^* = \frac{m}{1-\frac{mk_F}{3\pi^2}f_1^s} \tag{2.32}$$

which evidences that the effective mass can diverge for particularly strong interactions. This possibility was first introduced by Neville Mott [19], for a phenomenon now called Mott transition, where fermions localize. Such a transition occurs at zero temperature, in absence of thermal fluctuations, so it is an example of a quantum phase transition.
For the polarized unitary Fermi gas, which is the system of interest to our work, a recent work [1] predicted that the effective masses of the two components of the Fermi gas should diverge at a critical polarization, corresponding to a quantum phase transition to a Fulde-Farrel Larkin-Ovchinninkov (FFLO) phase superfluid phase [2][3]. In this phase Cooper pairs of spin up and spin down fermions condense in a state with a *finite momentum* of the pairs, differently from standard BCS superfluid with pairing at zero center-of-mass momentum of the pairs.
This prediction was obtained with approximate diagrammatic methods and requires an independent verification with alternative methods, such as QMC method that we will use in the present thesis's work.

2.4 Green's functions

Although Landau theory incorporates elements of a phenomenological theory, it has a solid microscopic foundation, in fact it is possible to specify the parameters in terms of microscopic scattering amplitudes of particles on the Fermi surface. Moreover, they have a connection to experimental measurable quantities. However, in order to inspect this theory microscopically we need a method to overcome the perturbation theory, since in general physical systems are non weakly interacting. This is based on Green's function formalism.
We can begin by considering a system of fermions which do not interact when the time $t \to -\infty$. The process of adiabatically switching on interactions can be thought of as a unitary transformation that acts on the state of a non-interacting system. As previously stated, a fundamental assumption in Fermi liquid theory is that there is a one-to-one correspondence between the physical particles of the system and the long-lived quasiparticle excitations near the Fermi surface. As a result, the states of both systems can be linked as follows:

$$\begin{aligned} |\Psi\rangle &= U|\Psi_0\rangle \\ |\overline{\mathbf{k}\sigma}\rangle &= U|\mathbf{k}\sigma\rangle \end{aligned} \tag{2.33}$$

where the first expression connects the ground states, while the second one connects the state where one fermion with momentum $\mathbf{k}$, close to Fermi surface, is added in the non interacting system to its adiabatic evolution at $t = 0$. The operator U can be written as time ordered exponential [17]

$$U = T\left[\exp\left\{-i\int_{-\infty}^{0} V(t)dt\right\}\right] \tag{2.34}$$

V represents the interaction in interaction representation. As t increases, we gradually turn on the interactions between the particles, until they reach full strength at $t = 0$. Since, in second quantization formalism, adding one fermion can be written as $|\mathbf{k}\sigma\rangle = c^\dagger_{\mathbf{k}\sigma}|\Psi_0\rangle$, the previous one-to-one correspondence becomes

$$|\overline{\mathbf{k}\sigma}\rangle = U c^\dagger_{\mathbf{k}\sigma} U^\dagger |\Psi\rangle \tag{2.35}$$

and we can identify the quasiparticle creation operator as

$$\overline{c}^\dagger_{\mathbf{k}\sigma} = U c^\dagger_{\mathbf{k}\sigma} U^\dagger \tag{2.36}$$

The excited single-particle states in the interacting system are occupied by quasi particles. Although $\overline{c}^\dagger_{\mathbf{k}\sigma}|\Psi\rangle$ carries the same momentum as the quasiparticle state $|\overline{\mathbf{k}\sigma}\rangle$, it may not be an energy eigenstate. It may be a superposition of states with different number of quasiparticle and quasihole pairs, all of these states can carry the same momentum but different energies. When they are sufficiently close to the Fermi surface they can closely approximate true excited eigenstates for long periods of time.
Even if the interactions between low lying quasi particles are weak, we cannot totally neglect them; interactions can be treated within a self-consistent field description. Landau's theory requires that the state created by applying the operator U to $|\mathbf{k}\sigma\rangle$ has a finite overlap, in the thermodynamic limit, with the state that one obtains by adding one bare particle to the initial ground state.
Now in order to proceed with a microscopic description, it is useful to introduce also the field operator as in [11]

$$\hat{\psi}_\alpha(\mathbf{r}) = \sum_\alpha \varphi_\alpha(\mathbf{r}) c_\alpha \tag{2.37}$$

where $\varphi(\mathbf{r})$ is a complete set of eigenfunctions of the single-particle Hilbert. In term of fields the number of particle operator reads

$$\hat{N} = \int \hat{\psi}^\dagger_\alpha(\mathbf{r})\hat{\psi}_\alpha(\mathbf{r}) d\mathbf{r} \ . \tag{2.38}$$

It is convenient to consider the Heisenberg representation for these operators, introducing an explicitly time dependence as in [11], then we write

$$\hat{\Psi}(t, \mathbf{r}) = e^{i\hat{K}t}\hat{\psi}_\alpha(\mathbf{r})e^{-i\hat{K}t} \tag{2.39}$$

where we set

$$\hat{K} = \hat{H} - \mu\hat{N} \ . \tag{2.40}$$

In terms of the field operators, one then defines the Green's function of a macroscopic system as

$$G_{\alpha\beta}(X_1, X_2) = -i\langle T[\hat{\Psi}_\alpha(X_1)\hat{\Psi}^\dagger_\beta(X_2)]\rangle \tag{2.41}$$

where time and spatial variables are represented by X, $\langle\cdot\rangle$ is an average over the ground state of the system, and T stands for a chronological ordered product.
For time homogeneity, and in case of spatially homogeneous system, one has

$$G_{\alpha\beta}(X_1, X_2) = G_{\alpha\beta}(X_1 - X_2) = G_{\alpha\beta}(X) \tag{2.42}$$

moreover in case of spin independent interactions

$$G_{\alpha\beta}(X_1, X_2) = G_{\alpha\beta}(X) = \delta_{\alpha\beta} G(X) \tag{2.43}$$

We can compute the momentum distribution of particles from a Fourier transformation of G

$$n(\mathbf{p}) = -i \int G(t, \mathbf{r})|_{t=0^-} e^{-i\mathbf{p}\cdot\mathbf{r}} d\mathbf{r} \tag{2.44}$$

It is convenient to express the Green's function in energy and momentum space instead of time and coordinate, these two representations are connected by

$$\begin{aligned} G(t, \mathbf{r}) &= \int G(\omega, \mathbf{p}) e^{i(\mathbf{p}\cdot\mathbf{r} - \omega t)} \frac{d\omega d\mathbf{p}}{(2\pi)^4} \\ G(\omega, \mathbf{p}) &= \int G(t, \mathbf{r}) e^{-i(\mathbf{p}\cdot\mathbf{r} - \omega t)} dt d\mathbf{r} \end{aligned} \tag{2.45}$$

We can also write the Green's function in a form called Lehmann representation

$$\begin{aligned} G_{\alpha\beta}(\omega, \mathbf{p}) &= \sum_m \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(0) | m, \mathbf{p} \rangle \langle m, \mathbf{p} | \hat{\psi}^\dagger_\beta(0) | \Psi_0 \rangle}{\omega - \omega_{m0} + i0^+} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(0) | m, -\mathbf{p} \rangle \langle m, -\mathbf{p} | \hat{\psi}_\alpha(0) | \Psi_0 \rangle}{\omega + \omega_{m0} - i0^+} \right] \\ &= \sum_m \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(0) | m, \mathbf{p} \rangle \langle m, \mathbf{p} | \hat{\psi}^\dagger_\beta(0) | \Psi_0 \rangle}{\omega - \xi^{(+)}_m + i0^+} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(0) | m, -\mathbf{p} \rangle \langle m, -\mathbf{p} | \hat{\psi}_\alpha(0) | \Psi_0 \rangle}{\omega + \xi^{(-)}_m - i0^+} \right] \end{aligned} \tag{2.46}$$

The Green's function in momentum representation is a meromorphic function of ω. In eq.(2.46) we expressed the ground state as $|\Psi_0\rangle$, and we introduced one set of states $|m, \mathbf{P}_m\rangle$ where the system has a number of particles $N_m = N_0 \pm 1$, and only the contribution for momentum $\mathbf{P}_m = \pm\mathbf{p}$ survives. The corresponding energies are

$$\begin{aligned} \omega_{m0} &= E_m - E_0 - \mu(N_m - N_0) = \xi^{(+)}_m \\ \omega_{0m} &= E_0 - E_m - \mu(N_0 - N_m) = \xi^{(-)}_m \end{aligned} \tag{2.47}$$

Then we have single poles at the exact excitation energies of the interacting system corresponding to a momentum $\mathbf{p}$. The singularities of Green's function immediately yield the energies of those excited states for which the numerator does not vanish. For an interacting system, the field operator connects the ground state with very many excited states of the system containing $N \pm 1$ particles. For the non-interacting system, however, the field operator connects only one state to the ground state, so that $G^0(\omega, \mathbf{p})$ has only a single pole. Discrete excitation energies ξ_m correspond to quasiparticle energies in Landau theory, each one corresponding to a specific momentum $\mathbf{P}_m$ of the system. This momentum can be associated to a single quasiparticle with a dispersion law, that corresponds to a singular point of Green's function.

In the thermodynamic limit the energy difference between levels becomes negligible with respect to the resolution, so we are allowed to replace the sum with an integral over a continuum of states. Assuming further spin independence, we can write

$$G(\omega, \mathbf{p}) = \sum_\alpha \frac{1}{2s+1} G_{\alpha\alpha}(\omega, \mathbf{p}) = \int d\omega' \left[\frac{A(\mathbf{p}, \omega')}{\omega - \omega' + i0^+} + \frac{B(\mathbf{p}, \omega')}{\omega + \omega' - i0^+} \right] \tag{2.48}$$

where we have introduced two spectral weight functions

$$\begin{aligned} A(\mathbf{p}, \omega) &= \sum_m |\langle m, \mathbf{p} | \hat{\psi}^\dagger_\alpha(0) | 0 \rangle|^2 \delta(\omega - \xi^{(+)}_m) \\ B(\mathbf{p}, \omega) &= \sum_m |\langle m, -\mathbf{p} | \hat{\psi}_\alpha(0) | 0 \rangle|^2 \delta(\omega - \xi^{(-)}_m) \end{aligned} \tag{2.49}$$

Then when $\omega < 0$ we can exploit the relation $A(\mathbf{p}, \omega) = B(\mathbf{p}, -|\omega|)$ and write the integral over the whole ω range

$$G(\omega, \mathbf{p}) = \int_{-\infty}^{\infty} \frac{d\omega' A(\mathbf{p}, \omega')}{\omega - \omega' + i0^+ \cdot \text{sign}(\omega')} \tag{2.50}$$

In an ideal Fermi gas we know exactly that the states obtained by adding a bare particle to the ground state are exact eigenstates. The excited state $\hat{\psi}_\alpha^\dagger(0)|0\rangle$ instead, is not in general an eigenstate of the interacting system, being a superposition of some eigenstates of different energies. For the ideal Fermi gas we know explicitly the spectrum, and

$$\begin{aligned} \xi_{\mathbf{p}}^{(+)} &= \frac{\mathbf{p}^2}{2m} - \frac{k_F^2}{2m} = \frac{\mathbf{p}^2}{2m} - \mu \\ \xi_{\mathbf{p}}^{(-)} &= \frac{k_F^2}{2m} - \frac{\mathbf{p}^2}{2m} = \mu - \frac{\mathbf{p}^2}{2m} \end{aligned} \tag{2.51}$$

Now we can compute the distribution of particles from Green's function as

$$n^{(0)}(\mathbf{p}) = -i \lim_{t \to 0^-} \int_{-\infty}^{\infty} G^{(0)}(\omega, \mathbf{p}) e^{-i\omega t} \frac{d\omega}{2\pi} = -\frac{i}{2\pi} \int \frac{d\omega}{\omega - \mathbf{p}^2/2m + \mu + i0^+ \cdot \text{sign}(p - k_F)} \tag{2.52}$$

for negative $t \neq 0$, we close the integration on real ω axis with a semi-circumference in upper half plane (then we can set $t = 0$). The integral assumes the value of the pole, that is absent in case of $p > k_F$, and is equal to one if $p < k_F$. So we recover the expected result for the ideal Fermi gas.

If we consider now the Fermi liquid, since we are assuming that $\hat{\psi}_\alpha^\dagger(0)|0\rangle$ is similar to an eigenstate for $\hat{H}$, we can associate an energy form similar to the non interacting case. Due to interactions the state changes its energy but can still be seen as a localized particle excitation, a dressed particle with a dressed energy. We expect a more complex form of the weight function $A(\mathbf{p}, \omega)$, which correspond to the probability density at energy ω for the state $\hat{\psi}_\alpha^\dagger(0)|0\rangle$. If it is an exact eigenstate with energy $\xi(\mathbf{p})$ then $A(\mathbf{p}, \omega)$ would simply be the delta function $\delta(\omega - \xi(\mathbf{p}))$.

Because of interactions we expect that the quasiparticle decays with a lifetime $1/\Gamma_{\mathbf{p}}$ then there will be a peak in $A(\mathbf{p}, \omega)$ which is centred on $\omega = \xi(\mathbf{p}) > 0$ and with a half-width of $\Gamma_{\mathbf{p}}$. The area under the peak will be a spectral strength $Z_{\mathbf{p}}$. The remaining contributions from states of other energies will make up a smoothly varying background. The smaller $Z_{\mathbf{p}}$ is in the range $0 \leq Z_{\mathbf{p}} \leq 1$, the greater will be the tendency of the single-particle excited states to mix with states involving excitations of two or more particles. For $Z_{\mathbf{p}} = 1$ there would be no mixing and the single-particle states would be pure eigenstates. For $Z_{\mathbf{p}} = 0$ the mixing with excitations of two or more particles would be complete leaving behind no detectable quasiparticle components. We can write the spectral function $A(\mathbf{p}, \omega)$ as a sum over excited states $|\boldsymbol{\lambda}\rangle$

$$A(\mathbf{p}, \omega) = \sum_{\boldsymbol{\lambda}} |M_{\boldsymbol{\lambda}}|^2 \delta(\omega - \xi_{\boldsymbol{\lambda}}) = Z_{\mathbf{p}} \delta(\omega - \xi_{\mathbf{p}}) + \sum_{\boldsymbol{\lambda} \neq \mathbf{p}} |M_{\boldsymbol{\lambda}}|^2 \delta(\omega - \xi_{\boldsymbol{\lambda}}) \tag{2.53}$$

where $|M_{\boldsymbol{\lambda}}|^2 = |\langle \boldsymbol{\lambda} | \hat{\psi}_\alpha^\dagger(0)|0\rangle|^2$, the first term in last equation indicates a sharp quasiparticle peak that corresponds to the pole of the Green's function. The second term is the difference with respect to a non interacting Fermi gas, a continuum of energy states. This peak changes with momentum, and gets narrow for $p \sim k_F$. When $p = k_F$ it becomes a delta function with weight Z. Depending on whether p is greater or smaller than k_F, the peak is for positive or negative frequency. In this latter case, it is actually the function $B(\mathbf{p}, \omega)$ which needs to be considered.

We can then compute the Green's function and use it to obtain the momentum distribution, with the approximation

$$\epsilon(\mathbf{p}) - \mu \approx v_F \left(p - k_F\right) \tag{2.54}$$

where now $v_F = k_F/m^*$, we can assume the following Green's function

$$G(\omega, \mathbf{p}) = \frac{Z_{\mathbf{p}}}{\omega - v_F \left(p - k_F\right) + i0 \cdot sign(\omega)} + g(\omega, \mathbf{p}) \tag{2.55}$$

where $g(\omega, p)$ is a regular function. We can compute now from G the Fermi step:

$$\lim_{q \to 0^+} n(k_F - q) - n(k_F + q) \tag{2.56}$$

Since g is regular, it does not contribute in this limit, and only the finite term of eq.(2.55) is relevant. In this limit we can replace the sign of ω with that of $(p - k_F)$, and obtain

$$\lim_{q \to 0^+} n(k_F - q) - n(k_F + q) = -i \int_{-\infty}^{\infty} \left[\frac{Z}{\omega + v_F q - i0^+} - \frac{Z}{\omega - v_F q + i0^+} \right] \frac{d\omega}{2\pi} = Z \ . \tag{2.57}$$

In computing this integral the exponential factor is not necessary for convergence and was omitted, and, since $n(\mathbf{p}) \leq 1$, it follows that $0 < Z \leq 1$; $Z = 1$ only for the ideal Fermi gas. Like for the ideal Fermi gas, there is still a discontinuity at $p = k_F$, but now the magnitude of this jump is not one, and $n(\mathbf{p})$ is different from zero also for $p > k_F$.

Following, e.g. Mahan [20] quite generally, we can introduce a self-energy correction $\Sigma(\mathbf{p}, \omega)$ that absorbs the effect of interactions on the single-particle Green's function $G(\mathbf{p}, \omega)$. The Dyson equation gives the relation between G for the interacting system and the non-interacting G^0 in terms of $\Sigma(\mathbf{p}, \omega)$

$$G = G^0 + G^0 \Sigma G \tag{2.58}$$

Then we can introduce the retarded Green's function

$$G^{\alpha\beta}_{\rm ret}(\omega, \mathbf{p}) = \sum_m \left[\frac{\langle \Psi_0 | \hat{\psi}_\alpha(0) | m, \mathbf{p} \rangle \langle m, \mathbf{p} | \hat{\psi}^\dagger_\beta(0) | \Psi_0 \rangle}{\omega - \xi^{(+)}_m + i0^+} + \frac{\langle \Psi_0 | \hat{\psi}^\dagger_\beta(0) | m, -\mathbf{p} \rangle \langle m, -\mathbf{p} | \hat{\psi}_\alpha(0) | \Psi_0 \rangle}{\omega + \xi^{(-)}_m + i0^+} \right] \tag{2.59}$$

that also has a Dyson equation

$$G_{\rm ret}(\mathbf{p}, \omega) = \frac{1}{\omega + i0^+ - \xi_{\mathbf{p}} - \Sigma_{\rm ret}(\mathbf{p}, \omega)} \tag{2.60}$$

where we introduced also the retarded self-energy, that can be used to write an expression for the spectral weight function

$$A(\mathbf{p}, \omega) = \frac{-2 {\rm Im}\Sigma_{\rm ret}(\mathbf{p}, \omega)}{\left[\omega - \xi_{\mathbf{p}} - {\rm Re}\Sigma_{\rm ret}(\mathbf{p}, \omega)\right]^2 + \left[{\rm Im}\Sigma_{\rm ret}(\mathbf{p}, \omega)\right]^2} \tag{2.61}$$

where the ${\rm Im}\Sigma_{\rm ret} \geq 0$ for positive $A(\mathbf{p}, \omega)$. Note that from here on, the function $A(\mathbf{p}, \omega)$ incorporates also $B(\mathbf{p}, \omega)$ by extending $A(\mathbf{p}, \omega)$ to negative frequencies with the definition $A(\mathbf{p}, \omega) = B(\mathbf{p}, |\omega|)$ for $\omega < 0$. If we take the limit ${\rm Im}\Sigma_{\rm ret} \to 0^-$ then

$$A(\mathbf{p}, \omega) = 2\pi \delta \left(\omega - \xi_{\mathbf{p}} - {\rm Re}\Sigma_{\rm ret}(\mathbf{p}, \omega)\right) \ . \tag{2.62}$$

Now we can exploit the properties of δ-function

$$\delta\left[g(x)\right] = \frac{\delta(x - x_0)}{|g'(x_0)|} \tag{2.63}$$

to rewrite the spectral function as

$$A(\mathbf{p}, \omega) = 2\pi Z(\mathbf{p})\delta\left(\omega - E_{\mathbf{p}} + \mu\right) \tag{2.64}$$

where

$$E(\mathbf{p}) - \mu = \xi_{\mathbf{p}} + \mathrm{Re}\Sigma_{\mathrm{ret}}(\mathbf{p}, E(\mathbf{p}) - \mu) \tag{2.65}$$

$$Z(\mathbf{p}) = \frac{1}{\left|1 - \frac{\partial}{\partial\omega}\Sigma_{\mathrm{ret}}(\mathbf{p}, \omega)\right|_{\omega = E_{\mathbf{p}} - \mu}} \tag{2.66}$$

This is a renormalization factor that is confined to the range $[0, 1]$ because of the sum rule

$$\int \frac{d\omega}{2\pi} A(\mathbf{p}, \omega) = 1 \tag{2.67}$$

satisfied by the spectral function. $Z(\mathbf{p})$ corresponds to the area under the peak of $A(\mathbf{p}, \omega)$, so it is well defined around the Fermi surface, where for $p = k_F$ it expresses the magnitude of the step in momentum distribution.
Now we can also connect the self-energy to the effective mass. Close to k_F we may set

$$E_{\mathbf{p}} = E_F + \frac{k_F}{m^*}(p - k_F) + O(p - k_F) \tag{2.68}$$

such that (setting $\epsilon_{\mathbf{p}} = p^2/2m$):

$$\frac{m}{m^*} = \left.\frac{\partial E_{\mathbf{p}}}{\partial \epsilon_{\mathbf{p}}}\right|_{p = k_F} = \left.\frac{m}{k_F}\frac{\partial E_{\mathbf{p}}}{\partial p}\right|_{p = k_F} \tag{2.69}$$

and the derivative of eq.(2.65) leads to the following relation for the effective mass of fermions on the Fermi surface (where the self-energy is real):

$$\frac{m}{m^*} = \left.\frac{1 + m/p\left(\partial/\partial p\right)\Sigma(\mathbf{p}, \omega)}{1 - \left(\partial/\partial\omega\right)\Sigma(\mathbf{p}, \omega)}\right|_{p = k_F, \omega = 0} \tag{2.70}$$

2.5 Extension to polarized system

Until now it has been considered a system of fermions in which the spin populations were equal, in this case both spin species have the same Fermi surface. However, we are interested here in a new situation in which one of the two population has a lower number of particles. This imbalance reflects in a polarization of the system computed as

$$p = \frac{N^{\uparrow} - N^{\downarrow}}{N^{\uparrow} + N^{\downarrow}} \tag{2.71}$$

In momentum space, one can describe the system as having two Fermi surfaces of radii $k_F^{\uparrow}$ and $k_F^{\downarrow} \neq k_F^{\uparrow}$. Because of this imbalance one can generalize the previous equation for the effective mass and quasiparticle residue as dependent on the self-energies for different spin species

$$Z_{\sigma}(\mathbf{p}) = \frac{1}{\left|1 - \frac{\partial}{\partial\omega}\Sigma^{\sigma}_{\mathrm{ret}}(\mathbf{p}, \omega)\right|_{\omega = \varepsilon_{\mathbf{p}}}} \tag{2.72}$$

$$\frac{m}{m^*_\sigma} = \frac{1 + m/p\,(\partial/\partial p)\,\Sigma_\sigma(\mathbf{p},\omega)}{1 - (\partial/\partial\omega)\,\Sigma_\sigma(\mathbf{p},\omega)}\bigg|_{p=k^\sigma_F,\omega=0} \tag{2.73}$$

In particular it is interesting for future simulations to see that there could be a difference between the effective masses of the two spin species.
One can try to analyse how the interaction function of the Fermi liquid theory changes for a finite value of the polarization. An exhaustive treatment can be found in [21], moreover in [22] Quader and Bedel derived a model for the induced interaction for arbitrarily polarized Fermi systems. They computed the interaction function as limits of the vertex function. Similar calculations are performed by Schwenk et al. [23] for a quasiparticle model of nuclear-nuclear interactions.
Still in the field of nuclear physics, but without introducing the complications of interactions, Sjöberg [24] generalised the expression for the quasiparticle effective mass to asymmetric nuclear matter with the aim of studying liquid phase of a neutron star. We can show briefly his method replacing imbalance of neutrons and protons with spins up and down.
On the basis of the expressions for the balanced case, one can rewrite the relation for the Fermi momenta of both spin populations as

$$k^\sigma_F = \left(3\pi^2 \frac{N^\sigma}{V}\right)^{1/3} \tag{2.74}$$

Following Sjöberg [24], let us consider a small displacement $\mathbf{q}$ of the minority Fermi surface (spin down), then the consequent energy variation (per unit volume) can be written as a change in potential energy plus the increase in kinetic energy

$$\delta E = \delta V^\downarrow(\mathbf{q}) + \frac{n_\downarrow \mathbf{q}^2}{2m} \tag{2.75}$$

Sjöberg exploited the properties of isotropy and Galilean invariance of the system to extract the relations for both effective masses. These conditions require no change in potential energy when both Fermi surfaces are displaced at the same time

$$\delta V^\uparrow(\mathbf{q}) = \delta V^\downarrow(\mathbf{q}) \tag{2.76}$$

Then, in combination with an energy functional that includes now a spin population dependency, and also a new interaction function with a spin population dependency, it is possible to generalize the result of eq.(2.29) as

$$\begin{aligned} \frac{m^*_\uparrow}{m} &= 1 + \frac{1}{3}N^*_\uparrow(0)\left[f_1^{\uparrow\uparrow} + \left(k^\downarrow_F/k^\uparrow_F\right)^2 f_1^{\uparrow\downarrow}\right] = 1 + \frac{1}{3}F_1^\uparrow \\ \frac{m^*_\downarrow}{m} &= 1 + \frac{1}{3}N^*_\downarrow(0)\left[f_1^{\downarrow\downarrow} + \left(k^\uparrow_F/k^\downarrow_F\right)^2 f_1^{\uparrow\downarrow}\right] = 1 + \frac{1}{3}F_1^\downarrow \end{aligned} \tag{2.77}$$

Where

$$N^*_\sigma(0) = \frac{m^*_\sigma k^\sigma_F}{2\pi^2} \tag{2.78}$$

and $f^{\uparrow\downarrow}$ can be expressed with the same expansion in Legendre polynomials as the original f in eq.(2.22).
One can now recover the initial result for a non-polarized system, writing the interaction function as

$$\begin{aligned} f^{\uparrow\uparrow} &= f^{\downarrow\downarrow} = f^{(s)} + f^{(a)} \\ f^{\uparrow\downarrow} &= f^{\downarrow\uparrow} = f^{(s)} - f^{(a)} \end{aligned} \tag{2.79}$$

If these relations are substituted in eq.(2.77), one is able to turn back to the initial case (equal density populations) of eq.(2.29), where the Fermi wavevectors are equal $k_F^{\uparrow} = k_F^{\downarrow} = k_F$ and also equal are the densities of states $N_{\downarrow}^*(0) = N_{\uparrow}^*(0) = N^*(0)/2$.

2.6 Breakdown of Fermi liquid theory

The theory discussed thus far can describe a large number of Fermi systems in condensed matter physics, i.e. it works well for most metals at low-temperature. However, other systems cannot be explained by Fermi liquid theory, such as one-dimensional Luttinger liquids or superconductors.
The work by Bardeen, Cooper and Schrieffer [25], resulting in the BCS theory of superconductivity, shows that the small attractive interaction between electrons, mediated by phonons, can produce electron pairs (Cooper pairs). BCS theory was designed to explain superconductivity in metals, but it quickly became a description of a broader class of phenomena, as for the pairing phenomenon in nuclear systems. BCS theory applies to an unpolarized system of fermions, i.e. with same density of up and down spins. A few years after the appearance of the BCS theory, the question of what happens when the system is polarized arose. In a system of electrons one can modify the spin populations by applying a magnetic field. Superconductivity is generally suppressed by a magnetic field. This can happen in two ways, the more effective one is the orbital effects due to Lorentz force, the second one is the Zeeman effect and the alignment of the spins which tends to break Cooper pairs.
A redefinition of the order parameter of the BCS theory can lead to new superconducting phases which are stable also for a finite amount of density imbalance. In particular, Fulde-Ferell [2] and separately Larkin and Ovchinnikov [3] considered a spatial dependence of the order parameter, that corresponds to pairing with a finite center-of-mass momentum. This can stabilize the system because of the partial overlap of the Fermi surfaces for the two spin species, resulting in a phase called FFLO.
The prediction of the coexistence of magnetism and superconductivity generated significant experimental validation efforts. Focusing on the FFLO phase, the goal of enforcing superconductivity suppression by spin imbalance, rather than by orbital effects, motivated the research for systems where orbital effects are absent. The Pauli suppression mechanism could be the dominant one in system where the Maki parameter [26] is high (at least larger than 1.8), such that the orbital limiting field is much higher than the Pauli one. One can find a discussion about the most promising materials for experimental evidence of the FFLO phase in [4] [6]. The first paper focuses mostly on heavy fermion superconductors, while the second one discusses two-dimensional layered superconductors.
Another possibility is to take advantage of systems of neutral atoms at ultralow temperatures. In this neutral case one has superfluidity rather than superconductivity, and the imbalance can be set by the realization of different pseudospin species. Then, this is a particularly suitable setup to study how Zeeman effect can suppress superfluidity. One can find an analysis of the FFLO state for ultracold quantum gases in lattice and harmonic potentials in [27].
In this thesis, we will investigate the behaviour of an atomic gas at zero temperatures, as a function of the gas polarization, in a box. We will describe the physical system in next chapter, along with what we know from theoretical calculations of its phase diagram. We will set in a particularly interesting condition of interaction between its atomic constituents, namely what is called the unitarity regime. We specifically want to monitor the behaviour of the effective mass in order to detect the possible presence of a divergence around a critical value of the polarization. This originates from the presence of a quantum critical point in the phase diagram

of the gas, that disrupt the "adiabatic" assumption underpinning Fermi liquid theory.

CHAPTER 3

ULTRACOLD FERMI GASES

In this chapter we will describe the system of interest to our work. It is a gas of ultra-cold atoms, which is an ideal experimental platform for testing theories of superfluids and superconductors. In fact, in these gases, one can easily control both interaction, and spin populations, or even study an exotic situation where the to spin species have different (bare) mass. Furthermore, also lattice structure are possible, allowing for previously unthinkable experiments in condensed matter physics.
The simulations that will be described in this thesis concern a gas of atoms at ultra-cold temperatures (we will actually work at $T = 0$). This kind of systems become only recently available to experimental analysis since ultra-low temperatures require a sophisticated cooling techniques. Laser and evaporative cooling, and magnetic trap must be employed to reach the temperature regime of nano-kelvin. Only at the end of the last century the ultracold regime, in which the atomic de-Broglie wavelength exceeds the typical inter-particle distance, was achieved. Then, when Bose-Einstein condensation (BEC) in dilute ultracold gases was obtained [28] [29], a new era in physics began. A few years later, also degenerate atomic Fermi gases were obtained [30] [31]. Then a large amount of experiments to validate the theories for quantum many-body systems were performed. In particular it was possible to inspect the BEC-BCS crossover.
We want to discuss in this chapter the main properties of these gases, and how their atoms interact. A much more detailed analysis of this physical system can be found in [32]. We will see that particular conditions will lead to an universal behaviour of the system, allowing for some simplifications of the physics under simulations.

3.1 Laser cooling and main experimental techniques

Ultra-cold gases are suitable for manipulation to set experimental conditions tu study Quantum Many-Body phenomena. First of all it is possible to cool them down to enhance their quantum properties, that are evident only when De Broglie wavelength Λ is of the order of (or larger than) the average interatomic distance

$$\Lambda = \sqrt{\frac{2\pi\hbar^2}{mk_BT}} \gtrsim n^{-1/3} \ . \tag{3.1}$$

In a typical trapped alkali gas experiments the number of atoms is around $10^5 - 10^9$, and the central density is $n \sim 10^{12} - 10^{15}$ cm^{-3}, so the temperature has to be reduced to $T \sim 1 - 100$ nK [33]. This process starts with laser cooling the system through transfer of momentum from high energy atoms by interaction with photons [34]. It is possible to reach temperatures around the microkelvin in this manner, but only if interaction between the atoms and the surrounding apparatus is avoided. This is done by trapping the atoms in a magnetic trap, as first realized in [35]. The more energetic atoms can then be released by lowering the depth of the trap, allowing one to reach ultralow temperatures at the expense of density. The remaining atoms re-thermalize to a lower temperature by elastic collisions. This process of *evaporative cooling* of the trapped quantum gas was introduced by Hess [36], setting a gas of spin-polarized atomic hydrogen at temperature of 30 μK and density of 10^{14}cm^{-3} by evaporative cooling and magnetic compression. At such low temperatures, the gas is metastable toward solidification, but the time scale for this transition is so long that we can ignore it during the experiment. When one works under these conditions, the kinetic energy of the atoms is so low that s-wave scattering dominates the interactions between atoms.
Besides their extremely low temperature, there are other important characteristics of ultracold quantum gases that make them so interesting for experiments. They are very clean systems: there are essentially no impurities in the experiments, and disorder is typically negligibly small, but it can be introduced on purpose, e.g., as in [37] to observe Anderson localization. This marks an important difference with respect to usual condensed matter systems, where impurities and defects effectively reduces the possibility to study ideal behaviours. Moreover, ultracold gases allow for a high experimental control of interactions, which are easily tunable by Feshbach resonances [38]. This has allowed for the experimental inspection of the physics of the BCS-BEC crossover, as done for ^{6}Li but also for different atoms. One can see a review on this topic in [39]. A condition for the interactions of particular interest is the so called "unitary regime", in which the scattering length diverge. This regime has been the focus on most experimental and theoretical works in the field [40].
One can further simulate lattice model by preparing an optical lattice. It consists of counter-propagating laser beams that combine and interfere to form a periodic potential for atoms by Stark effect. Both two and three dimensional lattices can be prepared. This periodic arrangement reproduces the naturally occurring structure in solid crystals. The first realization of electromagnetically trapped neutral sodium atoms was obtained by Jessen et al. [41].
Finally, the atoms of the gas can be selected in suitable internal quantum states. Radiofrequency sweeps with an adjustable sweep rate can be exploited to create a variable spin mixture of the same atom [42]. Moreover, these states have a time scale for spin flip much larger than the experimental time, so atoms can be considered with fixed spin population. This opens the possibility to study polarized systems, selecting with high accuracy the spin imbalance of the mixture.
Let us now examine the gas's physical regime; to do so, we must first select the dimensionless parameters that characterise its properties and use them as a references. The length scales that characterize the gas under exam are:

- The de Broglie wavelength: $\Lambda \sim T^{-1/2}$
- The average interatomic distance: $n^{-1/3}$
- The interaction potential range: R
- The s-wave scattering length: a

In the quantum degenerate regime $\Lambda \gtrsim n^{-1/3}$ so we are left with the remaining lengths. In addition, the diluteness condition implies that $Rn^{1/3} \ll 1$ in all experiments, and the interaction can be regarded as short-ranged. If $R \ll |a|$, the behaviour of the system is unusual, meaning that the specific details of the interaction potential are irrelevant, and only the scattering length matters. Most of the experiments are performed in this **...** regime. Finally, we distinguish the regime $|a|n^{1/3} \ll 1$, which corresponds to a weakly interacting gas, where perturbative methods work, and the strongly interacting regime $|a|n^{1/3} \gg 1$, where non-perturbative methods should be considered. The unitary Fermi gas corresponds to the extreme case $|a|n^{1/3} = \infty$. We are primarily interested in the universal, quantum, strongly-interacting regime, where the scattering length diverges.
In the following section an analysis of the interaction between atoms will be presented, as well as a discussion of the necessary conditions to achieve this optimal setting. We remark that for bosonic atoms, Stenger et al. [43] observed a strong increase of inelastic collisions rate in a sodium Bose-Einstein condensate to increasing $|a|$, leading to an instability of the condensate undergoing a rapid collisional decay. These significant losses severely constrain experiments with strongly interacting Bose gases. For fermions, instead, Petrov et al. [44] demonstrated that the collisional relaxation of the weakly bound dimers to deep bound states is suppressed by Fermi statistics, showing that the inelastic collisional rate is proportional to a^{-s}, where s is close to three, while the elastic collisional rate is proportional to a^2.

3.2 A recap of quantum theory of scattering

We now summarize the main elements of quantum theory of scattering that will be related to describe Feshbach resonances, as well as the trial wave function used for Monte Carlo simulations. We are here interested in analysing the two-body interaction between the atoms constituting a fermionic gas.
The time-independent Schrödinger equation simplifies in two parts, one for the centre of mass and one for relative motion. Focusing on equal masses, the time-independent wave function reads

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\mathrm{R}}(\mathbf{R}) \cdot \psi_{\mathrm{r}}(\mathbf{r}) \tag{3.2}$$

where $\psi_{\mathrm{R}}(\mathbf{R})$ is a plane wave while $\psi_{\mathrm{r}}(\mathbf{r})$ is determined by

$$H_{\mathrm{r}}\psi_{\mathrm{r}}(\mathbf{r}) = \left(-\frac{1}{m}\nabla^2_{\mathbf{r}} + V(\mathbf{r})\right)\psi_r(\mathbf{r}) = E_{\mathrm{r}}\psi_{\mathrm{r}}(\mathbf{r}) \tag{3.3}$$

(from here to the end of this chapter we set for convenience $\hbar = 1$). This equation can be used to describe a stationary state for a continuous beam of particles scattered from $V(r)$ (assumed central). We will also assume that the potential is short-range and therefore negligible at great distances. Then, in the limit of large distances, the scattered wavefunction must tend towards the free particle one.

3.2.1 Scattering wavefunction

Let us follow Stoof [33] and assume a plane-wave initial state perturbed by the scattering potential $\hat{V}$. So starting from the relative Schrödinger equation, writing $\hat{H} = \hat{H}_0 + \hat{V}$, one has

$$(2E_{\mathbf{p}} - \hat{H}_0)|\psi_{\mathbf{p}}\rangle = \hat{V}|\psi_{\mathbf{p}}\rangle \tag{3.4}$$

where $E_{\mathbf{p}} = p^2/2m_{\mathrm{r}}$. we obtain the Lippmann-Schwinger equation, that describes the scattered state

$$|\psi_{\mathbf{p}}^{(+)}\rangle = |\mathbf{p}\rangle + \frac{1}{2E_{\mathbf{p}} - \hat{H}_0 + i0^+}\hat{V}|\psi_{\mathbf{p}}^{(+)}\rangle \tag{3.5}$$

where $|\mathbf{p}\rangle$ is the unchanged initial state in the case of zero potential, and the infinitesimal term $i0$ avoids the singularity, while the $(+)$ sign will allow to obtain a final outgoing wave. We obtain the wavefunction $\psi_{\mathbf{p}}^{(+)}(\mathbf{r})$ with:

$$\langle \mathbf{r}|\psi_{\mathbf{p}}^{(+)}\rangle = \langle \mathbf{r}|\mathbf{p}\rangle + \int d\mathbf{r}'\langle \mathbf{r}|\frac{1}{2E_{\mathbf{p}} - \hat{H}_0 + i0}|\mathbf{r}'\rangle\langle \mathbf{r}'|\hat{V}|\psi_{\mathbf{p}}^{(+)}\rangle \tag{3.6}$$

where we introduced a complete set of eigenstates $|\mathbf{r}'\rangle$. By introducing a further set for $|\mathbf{p}'\rangle$ and exploiting the relation for previous transformation we obtain

$$\langle \mathbf{r}|\psi_{\mathbf{p}}^{(+)}\rangle = \frac{1}{(2\pi)^{3/2}}\left[\exp(i\mathbf{k}\cdot\mathbf{r}) + \int d\mathbf{r}'\int d\mathbf{p}'\frac{\exp(i\mathbf{p}'\cdot(\mathbf{r}-\mathbf{r}'))}{2E_{\mathbf{p}} - 2E_{\mathbf{p}'} + i0}\langle \mathbf{r}'|\hat{V}|\psi_{\mathbf{p}}^{(+)}\rangle\right] , \tag{3.7}$$

yielding

$$\langle \mathbf{r}|\psi_{\mathbf{p}}^{(+)}\rangle = \frac{1}{(2\pi)^{3/2}}\left[\exp(i\mathbf{k}\cdot\mathbf{r}) + \int d\mathbf{r}'\frac{\exp(ip|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|}\langle \mathbf{r}'|\hat{V}|\psi_{\mathbf{p}}^{(+)}\rangle\right] \tag{3.8}$$

Since we are interested in large distances from scattering centre, $r \gg r'$, the modulus in the exponential can be approximated as $r\sqrt{1 - \frac{2rr'}{r^2} + \frac{r'^2}{r^2}} \approx r - \frac{\mathbf{r}}{r}\cdot\mathbf{r}'$, while at the denominator, because of much lesser sensibility, we can take just r.
If we define $\mathbf{p}' = \frac{p\mathbf{r}}{r}$, we obtain

$$\langle \mathbf{r}|\psi_{\mathbf{p}}^{(+)}\rangle = \psi_{\mathbf{p}}^{(+)}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}}\left[\exp\left(i\mathbf{p}\cdot\mathbf{r}\right) + f(\mathbf{p}',\mathbf{p})\frac{\exp\left(ipr\right)}{r}\right] \tag{3.9}$$

where the scattering amplitude $f(\mathbf{p}',\mathbf{p})$ is given by

$$f(\mathbf{p}',\mathbf{p}) = -(2\pi)^{3/2}\frac{m}{4\pi}\langle \mathbf{p}'|\hat{V}|\psi_{\mathbf{p}}^{(+)}\rangle \tag{3.10}$$

So, in conclusion, we obtain the superposition of an incident plane wave and a spherical outgoing wave.

Partial Wave Expansion

Replacing momenta with the corresponding wave vectors, the scattering wavefunction becomes

$$\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}}\left[\exp(i\mathbf{k}\cdot\mathbf{r}) + f(\mathbf{k}',\mathbf{k})\frac{\exp(ikr)}{r}\right] \tag{3.11}$$

If $V(\mathbf{r}_1,\mathbf{r}_2)$ is central, then the scattering amplitude will depend only on the wave vectors and the scattering angle,.The scattering amplitude and the incoming plane wave can then be expanded in partial waves as

$$\begin{aligned} f(\mathbf{k}'.\mathbf{k}) &= \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\cos\theta) \\ e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{l=0}^{\infty}(2l+1)j_l(kr)P_l(\cos\theta) \end{aligned} \tag{3.12}$$

where $P_l(\cos\theta)$ are Legendre polynomials and $j_l(x)$ Bessel spherical functions of first order. Considering again large distances, we can take the asymptotic limit of $j_l(x)$ for large x and write

$$e^{i\mathbf{k}\cdot\mathbf{r}} \approx \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\left(\frac{e^{ikr}-e^{-i(kr-l\pi)}}{2ikr}\right) \tag{3.13}$$

that leads to

$$\begin{aligned}\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) &= \frac{1}{(2\pi)^{3/2}}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\left[\left(\frac{e^{ikr}-e^{-i(kr-l\pi)}}{2ikr}\right)+f_l(k)\frac{e^{ikr}}{r}\right] \\ &= \frac{1}{(2\pi)^{3/2}}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\frac{1}{2ik}\left[-\frac{e^{-i(kr-l\pi)}}{r}+\left(1+2ikf_l(k)\right)\frac{e^{ikr}}{r}\right]\end{aligned} \tag{3.14}$$

At large distances, the effect of the collision process can be seen as change of the coefficient in the outgoing partial wave

$$\frac{e^{ikr}}{r} \longrightarrow \frac{\left(1+2ikf_l(k)\right)e^{ikr}}{r} \tag{3.15}$$

and, enforcing the conservation of flux probability by setting

$$1+2ikf_l(k) \equiv e^{2i\delta_l(k)} \tag{3.16}$$

such that $|1+2ikf_l(k)|=1$, one sees that all the information on the scattering process is know contained in the phase shift $\delta_l(k)$.

S-wave approximation

As anticipated before, for a dilute and ultracold gas, the typical wavevector $\mathbf{k}$ is small. Therefore, the dominant contribution is for $l=0$. In this case, the s-wave scattering length is defined as

$$a = -\lim_{k\to 0}\frac{\delta_0(k)}{k} \tag{3.17}$$

and the scattering amplitude

$$f(\mathbf{k},\mathbf{k}') \approx f_0(k) = \frac{e^{2i\delta_0(k)}-1}{2ik} \approx \frac{1}{k\cot(\delta_0(k))-ik} \approx -a \tag{3.18}$$

where in last step we performed the expansion for very low energies, or $\lim k \to 0$.

3.2.2 Spherical well or barrier

Let us move now to the analysis of a more specific cases related to this thesis: the scattering from a potential well, or barrier. We will follow [45], where one can find a complete analysis of this and other cases. The equation for the radial wavefunction $\psi_{\mathrm{r}}(\mathbf{r}) = u(\mathbf{r})/r$ reads

$$\frac{d^2}{dr^2}u(r) + 2m_r\left(E_{\mathrm{r}} - \frac{-l(l+1)}{2m_r r^2} - V(r)\right)u(r) = 0 \tag{3.19}$$

where $m_r = m/2$ for equal masses. By setting further

$$\begin{aligned}k^2 &= mE_r \\ U(r) &= mV(r)\end{aligned} \tag{3.20}$$

the previous equation can be written as

$$\frac{d^2}{dr^2}u(r) + \left(k^2 - \frac{l(l+1)}{r^2} - U(r)\right)u(r) = 0 \ . \tag{3.21}$$

If we consider the spatial region where (supposing $U(r)$ decays faster than $1/r^2$ at large r), setting $z = kr$ and $\rho(r) = u(r)/z$, the Schrödinger equation turns into the spherical Bessel equation

$$\frac{d^2\rho}{dz^2} + \frac{2}{z}\frac{d\rho}{dz} + \left(1 - \frac{l(l+1)}{z^2}\right)\rho(r) = 0 \tag{3.22}$$

which has known solutions. We can write the radial function as a combination of these solutions

$$u_{kl}(r) = A_{kl}krj_l(kr) + B_{kl}krn_l(kr) \tag{3.23}$$

where j_l and n_l are spherical Bessel function of first and second type. At large distances, the previous solution becomes

$$u_{kl}(r) \approx A_{kl}\sin(kr - l\pi/2) - B_{kl}\cos(kr - l\pi/2) \tag{3.24}$$

that can be normalized by rewriting

$$A_{kl}^2 + B_{kl}^2 = 4 \tag{3.25}$$

Then it is possible to write

$$A_{kl} = 2\cos(\delta_l(k)), \ \ B_{kl} = -2\sin(\delta_l(k)) \tag{3.26}$$

with

$$\cot(\delta_l(k)) = -\frac{A_{kl}}{B_{kl}} \tag{3.27}$$

where $\delta_l(k)$ is the phase shift. One then has at large r:

$$u_{kl}(r) \approx 2\sin(kr - l\pi/2 + \delta_l(k)) \tag{3.28}$$

For a spherical well/barrier:

$$V(r) = \begin{cases} V_0 \text{ if } r < R \\ 0 \text{ if } r \geq R \end{cases} \tag{3.29}$$

since for continuous potential the Schrödinger equation leads to continuity of $u(r)$ and its derivative, we must enforce

$$\left.\frac{u'(r)}{u(r)}\right|_{r=R^-} = \left.\frac{u'(r)}{u(r)}\right|_{r=R^+} \tag{3.30}$$

For $r > R$, eq.(3.23) and eq.(3.26)

$$u(r) = u_{kl}(r) = 2kr\left[\cos(\delta_l(k))j_l(kr) - \sin(\delta_l(k))n_l(kr)\right] \tag{3.31}$$

It is convenient to multiply both side of eq.(3.30) by r, then we have

$$\begin{aligned}
&\left.\frac{r}{u(r)}\frac{du_{kl}(r)}{dr}\right|_{r=R^+} = \\
&= \left.\frac{kr\left[\cos(\delta_l(k))j_l(kr) - \sin(\delta_l(k))n_l(kr)\right] + (kr)^2\left[\cos(\delta_l(k))j_l'(kr) - \sin(\delta_l(k))n_l'(kr)\right]}{kr\left[\cos(\delta_l(k))j_l(kr) - \sin(\delta_l(k))n_l(kr)\right]}\right|_{r=R^+} = \\
&= 1 + kr\left.\frac{\cos(\delta_l(k))j_l'(kr) - \sin(\delta_l(k))n_l'(kr)}{\left[\cos(\delta_l(k))j_l(kr) - \sin(\delta_l(k))n_l(kr)\right]}\right|_{r=R^+} \\
&= 1 + kR\frac{\cot(\delta_l(k))j_l'(kR) - n_l'(kR)}{\left[\cot(\delta_l(k))j_l(kR) - n_l(kR)\right]}
\end{aligned} \tag{3.32}$$

By defining now (for given k and l)

$$L_{kl} = \frac{r}{u(r)} \frac{du_{kl}(r)}{dr}\bigg|_{r=R^-} \tag{3.33}$$

and setting $\gamma = kR$, the continuity by eq.(3.30) yields

$$\cot(\delta_l(k)) = \frac{(L_{kl}-1)n_l(\gamma) - \gamma n_l'(\gamma)}{(L_{kl}-1)j_l(\gamma) - \gamma j_l'(\gamma)} \tag{3.34}$$

For $l = 0$, the analysis is particularly simple, the radial Schrödinger equation reads

$$\frac{d^2u(r)}{dr^2} + \left(k^2 - U(r)\right) u(r) = 0 \tag{3.35}$$

Defining a new wavevector as $\kappa^2 = k^2 - U$ the equation 3.35 becomes

$$\begin{aligned} \frac{d^2u(r)}{dr^2} + \kappa^2 u(r) = 0 \text{ if } r < R \\ \frac{d^2u(r)}{dr^2} + k^2 u(r) = 0 \text{ if } r \geq R \end{aligned} \tag{3.36}$$

The solution of this problem depends on the sign of V_0 and energy. We have the following possibilities:

- In the case $V_0 > 0$, for energies greater than the barrier height V_0, the solution is

$$\begin{aligned} u(r) = A\sin(\kappa r) \text{ if } r < R \\ u(r) = B\sin(kr + \delta_0(k)) \text{ if } r \geq R \end{aligned} \tag{3.37}$$

  and, after enforcing the continuity of $r = R$

$$\frac{u'(r)}{u(r)}\bigg|_{r=R^-} = \frac{u'(r)}{u(r)}\bigg|_{r=R^+} \tag{3.38}$$

  one gets

$$k\tan(\kappa R) = \kappa\tan(\delta_0(k) + kR) \tag{3.39}$$

  and so

$$\delta_0(k) = -kR + \tan^{-1}\left[\frac{k}{\kappa}\tan(\kappa R)\right] \tag{3.40}$$

- In the case $V_0 > 0$, for energies smaller than barrier height V_0, the solution is

$$\begin{aligned} u(r) = A\sinh(\kappa r) \text{ if } r < R \\ u(r) = B\sin(kr + \delta_0(k)) \text{ if } r \geq R \end{aligned} \tag{3.41}$$

  and the continuity in $r = R$ yields

$$k\tanh(\kappa R) = \kappa\tan(\delta_0(k) + kR) \tag{3.42}$$

  and so

$$\delta_0(k) = -kR + \tan^{-1}\left[\frac{k}{\kappa}\tanh(\kappa R)\right] \tag{3.43}$$

This is the relevant case to calculate the scattering length for the repulsive potential, hence

$$a = -\lim_{k\to 0} \frac{\delta_0(k)}{k} \tag{3.44}$$

and thus

$$a = R\left(1 - \frac{\tanh(\kappa R)}{\kappa R}\right) \tag{3.45}$$

Note that for $V_0 \to \infty$, we obtain the solution $a = R$.

- In the case $V_0 < 0$, the potential is attractive and bound states are possible. Now the solution of the Schrödinger equation for positive energies is equivalent to the case of positive potential with energies greater than the potential height V_0, while for negative energies ($V_0 < E < 0$)

$$\begin{aligned} u(r) &= A\sin(\kappa r) \text{ if } r < R \\ u(r) &= Be^{-rk'} \text{ if } r \geq R \end{aligned} \tag{3.46}$$

where, because of negative energy we defined $E = -(k')^2/m$. After enforcing continuity at $r = R$, we get the relation

$$k' = -\kappa\cot(\kappa R) \tag{3.47}$$

and for energies $|E|$ much smaller than $|V_0|$, eq.(3.47) may be written as

$$\sqrt{m|E|} = -\sqrt{m|V_0|}\cot\left(R\sqrt{m|V_0|}\right) = -\frac{\zeta}{R}\cot(\zeta) \tag{3.48}$$

where

$$\zeta = R\sqrt{m|V_0|}\ . \tag{3.49}$$

At the same time, in the attractive case, the scattering length is

$$a = R\left(1 - \frac{\tan(\zeta)}{\zeta}\right)\ . \tag{3.50}$$

Note now that $|E| \ll |V_0|$ when $\cot\zeta \to 0$ and thus for $\zeta \to \infty$. In this situation

$$a = -R\frac{\tan\zeta}{\zeta} \tag{3.51}$$

and then

$$-\frac{\zeta}{R}\cot(\zeta) = \frac{1}{a}\ . \tag{3.52}$$

We thus have that:

$$\sqrt{m|E|} \approx \frac{1}{a}\ , \tag{3.53}$$

that is

$$E \approx -\frac{1}{ma^2}\ . \tag{3.54}$$

Let us now summarize our results. We have seen that in the case of low energies, the scattering can be described in terms of a single parameter, the s-wave scattering length a. Then all potentials, that lead to the same scattering length a are equivalent. Therefore, it is possible to select the simplest and most convenient potential. When the potential is attractive, the s-wave

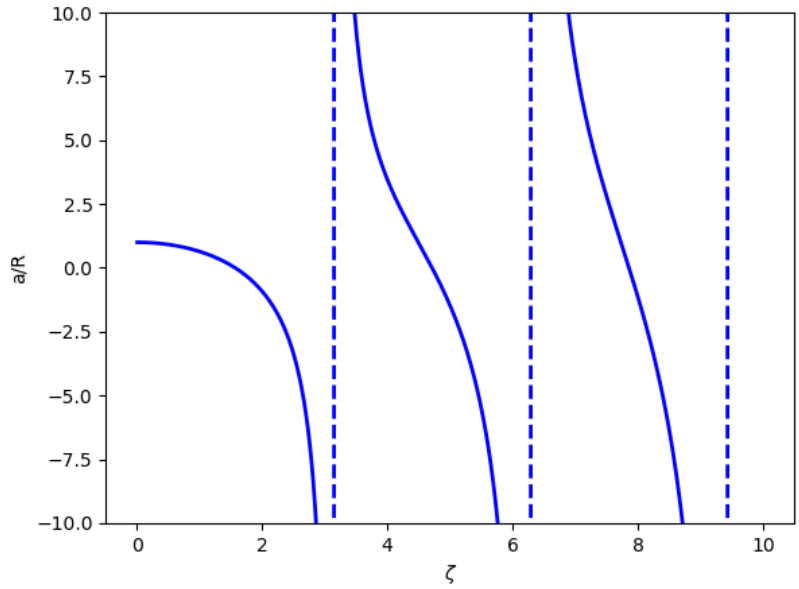


Figure 3.1: Plot of s-wave scattering length as function of $\zeta = R\sqrt{m|V_0|}$.

scattering length can be both positive or negative. It is zero when

$$R\sqrt{m|V_0|} = \tan\left(R\sqrt{m|V_0|}\right) \longleftrightarrow \zeta = \tan(\zeta) \tag{3.55}$$

so the scattering length has a recurrent behaviour as function of potential depth as in Fig.3.1. The divergence occurs with a sign change when it is fulfilled the following condition

$$\zeta = \left(n + \frac{1}{2}\right)\pi \tag{3.56}$$

When this occurs, there is a potential resonance, and the system is in the unitary regime, and the gas is at the same time dilute (because the range of the interatomic potential is much smaller than the inter-particle distance) and strongly interacting gas (because the scattering length is much larger than the inter-particle distance) [32]. The condition eq.(3.56) corresponds to the appearance of a new bound state in the two-body problem. When the potential strength is close to satisfy this condition, $|a|$ will be large and the scattering at low energy resonant.

3.2.3 Two-body transition matrix

In the previous subsection we discussed the wave function of the two body scattering problem. The same results can also be reached with another method, which is more convenient in describing further properties of ultra-cold fermionic gases. Still following [33], let us consider again eq.(3.5), but now it is useful to replace momentum $\mathbf{p}$ with wavevector $\mathbf{k}$, and multiply both side by $\hat{V}$. The Lippmann-Schwinger equation can then be reformulated by introducing the 2-body transition matrix $\hat{T}^{2B}$, such that

$$\hat{T}^{2B}(z)|\mathbf{k}\rangle = \hat{V}|\mathbf{k}\rangle + \hat{V}\frac{1}{z - \hat{H}_0}\hat{T}^{2B}(z)|\mathbf{k}\rangle \tag{3.57}$$

where z is, in general, a complex energy. Eq.(3.57) must hold for any $|\mathbf{k}\rangle$, so $\hat{T}^{2B}$ must satisfy the equation

$$\hat{T}^{2B}(z) = \hat{V} + \hat{V}\frac{1}{z - \hat{H}_0}\hat{T}^{2B}(z) \tag{3.58}$$

The previous equation can be solved by iteration, leading the Born series:

$$\hat{T}^{2B}(z) = \hat{V} + \hat{V}\hat{G}_0(z)\hat{V} + \hat{V}\hat{G}_0(z)\hat{V}\hat{G}_0(z)\hat{V} + ... \tag{3.59}$$

where

$$\hat{G}_0(z) = \frac{1}{z - \hat{H}_0} \tag{3.60}$$

The Born approximation corresponds to retain just the first term. The formal solution is instead given by

$$\hat{T}^{2B}(z) = \hat{V} + \hat{V}\frac{1}{z - \hat{H}}\hat{V} = \hat{V} + \hat{V}\hat{G}(z)\hat{V} \tag{3.61}$$

We can see from the last equation that the two-body transition matrix has singularities in correspondence with the exact eigenenergies of two-body problem. In terms of $\hat{T}^{2B}$ the scattering amplitude becomes

$$f(\mathbf{k}', \mathbf{k}) = -(2\pi)^{3/2}\frac{m}{4\pi}\langle \mathbf{k}'|\hat{T}^{2B}(z)|\mathbf{k}\rangle \tag{3.62}$$

At low energies:

$$f(\mathbf{k}', \mathbf{k}\ ) \approx f_0(k) = \frac{1}{k\cot(\delta_0(k)) - ik} \tag{3.63}$$

and

$$k\cot(\delta_0(k)) \approx -\frac{1}{a} + \frac{1}{2}r_{\text{eff}}k^2 + ... \tag{3.64}$$

Then the T^{2B} matrix can be approximated as

$$\langle \mathbf{k}'|\hat{T}^{2B}(z)|\mathbf{k}\rangle \approx -\frac{4\pi}{m}f_0(k) = -\frac{4\pi}{m}\frac{1}{k\cot(\delta_0(k)) - ik} \approx \frac{4\pi a}{m} \tag{3.65}$$

3.3 Feshbach resonances

As already highlighted above, the scattering length plays a very important role in describing the ultracold quantum gases. In fact it can be interpreted as giving the effective interaction strength. Inouye et al. [46] analysed the scattering for sodium atoms, and showed that the interaction strength can be varied in a wide range by means of a magnetic field. This effect is now known as Feshbach resonance [38], and it leads to the possibility of studying how a system behaves when its interaction changes. This phenomenon is much more flexible than the case of potential resonance, where one needs to change of shape of the potential to obtain a resonance. We will see that in this case one has a high degree of control on the energy difference tunable by a magnetic field. Let us study this phenomenon by means of a two-channel model following [47].

3.3.1 Two-channel scattering

We are now considering a scenario in which two incoming atoms of variable energies collide. They are in two different states, i.e., the atom's lowest two hyperfine states. So they have some magnetic moment that is different from the magnetic moment of a possible bound state,

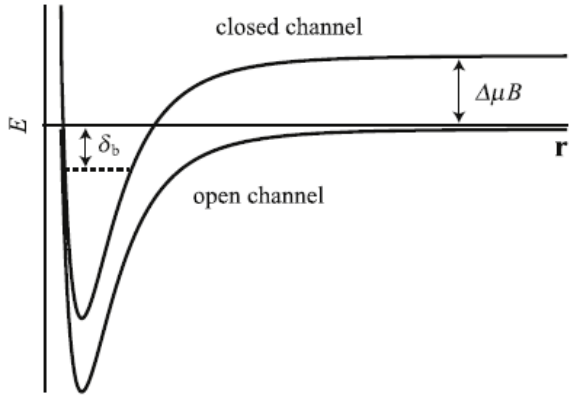


Figure 3.2: Illustration of the two-channel structure of a Feshbach resonance from [47] . The atomic (open) channel and the molecular (closed) channel are shown in the absence of a coupling between them. The interaction potential in the open channel is called the background interaction $V_{bg}(r)$, and is function of the interatomic separation r. The bound state in the closed channel is called the bare molecular state. It has a small energy difference δ_b with the atomic continuum. The Zeeman shift $\Delta\mu B$ separates the two channels.

and there may be some energy difference due to the Zeeman effect. This energy, which can be controlled by varying the magnetic field, is essential in determining which of two scattering channels is open and which is closed. When two atoms collide, they can form a molecule that rapidly decays (it is a virtual molecule) when the atoms are far apart if the collision energy is low relative to the Zeeman energy difference. Then we say that the atomic channel is open, while the molecular channel is closed.

If by now, we neglect coupling between channels, we can describe the molecule in terms of bare quantities, the wavefunction $|\psi_m\rangle$ and energy δ_b, or bare detuning, a small energy difference with respect to atomic continuum, as in figure 3.2. It is a solutions of the Schrödinger equation in the closed channel, decoupled from the open one. In the open channel, the atoms interact through the background potential $\hat{V}_{\rm bg}$.

Let us now allow the channels to be coupled by a potential $\hat{V}_{\rm am}$. It arises from the possibility of spin flip because of hyperfine interaction. We have to now to replace bare quantities with dressed ones for coupled channel, so we call the dressed "molecular" state $|\psi_{\rm dr}\rangle$ with energy E. If we call B_0 the field corresponding to the resonance, the energy shift becomes $\delta = \Delta\mu(B-B_0)$. It is the energy shift of the bound state with respect to the continuum of scattered states, that is also called the experimental detuning.

We can define the total interaction that acts in the open channel as $\hat{V} = \hat{V}_{\rm bg} + \hat{V}_m$, where $\hat{V}_m$ is the molecule-mediated interaction. This is the potential we have to consider in computing the transition matrix in the open channel. One finds that it consists of two terms, one for the single background plus one given by the resonance. In the limit of zero energy, the first leads to the same relation as before with s-wave scattering length $a_{\rm bg}$ for the background interaction, while for the second one has [47]: $T^{2B}_{res}(0) = -g^2/\delta$, that is divergent at resonance. The total T-matrix is

$$T^{2B}(0) = T^{2B}_{bg}(0) + T^{2B}_{res}(0) = \frac{4\pi^2 a_{bg}}{m} - \frac{g^2}{\delta} = \frac{4\pi^2 a_{bg}}{m}\left[1 - \frac{\Delta B}{B - B_0}\right] , \tag{3.66}$$

where we expressed g ain terms of experimental parameters

$$g = \sqrt{4\pi^2 a_{bg}\Delta B \Delta\mu/m} \tag{3.67}$$

and ΔB defines the width of the resonance. We have thus obtained a simple expression for s-wave scattering length in the presence of a Feshbach resonance as function of the magnetic field B. We recover the previous concept of unitarity for a proper value $B = B_0$, since it leads to the divergence of the scattering length.
We have to underline the distinction between broad and narrow resonances, which is based on the comparison between the Fermi wavevector k_F and the effective interaction range R. One has a broad resonance when $k_F|R| \ll 1$. In this case, the effective range loses importance at the many-body level and one can use $k_F a$ as the only parameter of interaction. On the contrary, for narrow resonance, the effective range is negative and becomes a relevant scale of the problem [32].
In addition, it is possible to extract information on the dressed molecule. For example, for the broad resonance of ^{6}Li (at about 834 G) it has been found [48] [49] that the dressed molecule has a very small amplitude in the closed channel over the resonance. In this range we can consider only the open channel, and the bare molecular state just allows the resonance. Then, we can describe the many body system with a single channel Hamiltonian with the transition matrix in eq.(3.66). However, with respect to a single uncoupled channel, the coupling interaction leads to a shift in the bound state energy for the resonance, as well as the field value at resonance. Furthermore, the coupling between channels switches the bound-state energy's dependence on the magnetic field from linear to quadratic near the resonance [33].

3.3.2 Feshbach resonance for ^{6}Li

Let us consider the case of ^{6}Li atoms as an example of what we discussed in previous subsection. It is an alkali atom, so one can approximate it as a core of charge e that interact by Coulomb interaction with a single electron of charge $-e$. The solution of the corresponding Schrödinger equation leads to a degenerate set of eigenenergies. This degeneracy can be lifted by introducing electron-nuclear spin coupling corrections, the hyper-fine structure, and application of magnetic field (while the spin-orbit coupling does not lead to fine-structure splitting for alkali atom in the ground state since $L = 0$). For zero angular orbital momentum, the total angular quantum number f, its projection on the quantization axis m_f, and the nuclear and electronic spin quantum numbers i and s, completely determine the hyperfine spin state. We can write the eigenstate for the s-orbital ground state of an alkali atom as $|nl; (is)fm_f\rangle = |00; (is)fm_f\rangle$, that splits in two manifold of degenerate states for $f = i \pm s$. This degeneracy can be lifted by the application of an external field. Depending on the strength of the field, the Zeeman interaction could be a perturbation of the hyperfine interaction (weak field), where the splitting is defined by $|fm_f\rangle$, or it could be the fundamental interaction perturbed by the hyperfine one (strong field). In this second case the splitting is defined by $|im_i, sm_s\rangle$. Some of the eigenstates have to bend in energy in order to connect smoothly these two limits.
We can consider a two-component mixture prepared with the two lowest hyperfine states denoted by $|1\rangle$, $|2\rangle$ in figure 3.3b. The composition of the system for the experiments is chosen using a radio frequency sweep with an adjustable sweep rate, a technique able to control independently the population of the two lowest hyperfine states of ^{6}Li allowing for the study of polarized Fermi system. A very broad Feshbach resonance occurs generated for $B_0 = 834$ G as shown in figure 3.3a.

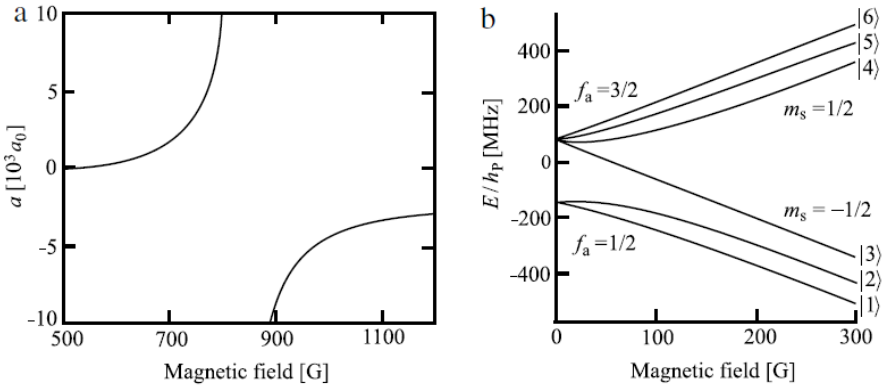


Figure 3.3: a) The scattering length in units of the Bohr radius a_0 as a function of magnetic field strength, the system is a gas of ^{6}Li atoms that has a very broad resonance for 834G. (b) Hyperfine structure of the electronic ground state in ^{6}Li as a function of magnetic-field strength, where E is the internal energy and h_P is Planck's constant. The total angular momentum of ^{6}Li if $f_a = 1 \pm 1/2$. Figure from [47].

3.4 Polarized Fermi gas

So far, we have discussed the general properties of ultracold quantum gases and how their atomic constituents interact. Then, in such a physical system, we described how to tune the interaction strength. Now, we wish to discuss what happens to such a gas when the two spin populations are imbalanced.
Similar to what was discussed in [1], we will consider a spin-1/2 Fermi gas with an attractive interaction, where starting from limit of high polarization, we will continuously reduce the population imbalance of the system. We will be interested in particular in the unitary limit of the interaction, corresponding to a diverging scattering length. We expect that for large polarization the system is normal (non-superfluid) even at zero temperature. At such large polarization, the system can be connected, in a continuous way to a non-interacting polarized Fermi gas, and is thus expected [50] [51] to behave as a Fermi liquid with two kind of quasi-particles (corresponding to the two spins) and a different behaviour depending on the polarization. The limit of large polarization $p \sim 1$ is generally referred to as the Fermi polaron limit. In the next subsection we will discuss this limit, which has drown a great interest in the recent literature. After that, we will present the theoretical prediction of what could happen with decreasing the polarization.

3.4.1 Polaron problem

The term polaron was introduced by Landau [52] and Pekar [53], to describe the properties of conduction electrons in dielectric materials: it is a quasi-particle formed by the dressing of the electrons by collective excitations in the material. This idea was later elaborated in a quantum mechanical theory by Fröhlich [54]. The same term appears also in the physics of ultracold gas, but here refers o what is obtained through population difference in the gas, with the minority atoms playing the role of impurities, while the majority atoms play the role of the medium. In this case we can describe the excitations as Fermi polarons, which are realizations of the quasiparticle concept.
The investigation of this low-temperature atomic system revealed that many properties of the polaron can be described in the strong coupling regime using a relatively simple theory. Since

accurate theories for strongly interacting systems are rare in many-body physics, the polaron problem is an important test for understanding strongly interacting systems. The study of a single impurity, in particular, allows for accurate information on strongly interacting polarized gases, even when the minority particle concentration is finite. Fermi liquid theory describes the behaviour of these systems.
Note that for a single spin down fermion in a large bath of spin up fermions, the latter can be treated as an ideal Fermi gas (we recall that interactions between identical fermions are highly suppressed in dilute ultracold Fermi gases).

Simple Fermi polaron model

In [55] Massignan et al. analysed first one simple model where the impurity interacts only with the neighbouring majority atoms, while the presence of atoms in the Fermi sea is taken into account as a Fermi pressure acting on spin-↑ particles. This is done by solving the Schrödinger equation for a pair of atoms $(\uparrow,\downarrow)$, namely a 2-body problem, and requiring that the wave function

$$\psi = \frac{1}{r}\sin(kr+\delta) \tag{3.68}$$

satisfies the boundary condition $\delta = -kR$, so it vanishes at the boundary of a spherical cavity of radius R (one can find this radius by equating the ground state energy in absence of interaction with the energy needed to add a pair of $(\uparrow,\downarrow)$ particles to the Fermi sea).
This model leads to qualitative description of the evolution of the polaron when changing the interaction strength. The weak interaction limit is correctly represented, but the model loses its effectiveness as the interaction becomes stronger. For scattering length $a > 0$, the attractive polaron approaches the energy of the dimer in the vacuum, this because the ground state in this case is represented by a strongly bound dimer immersed in the Fermi sea.
This simplified model is able to show the presence of a repulsive branch of excitations, however, in the vicinity of the resonance is unable to describe the metastability of the repulsive branch, and also the presence of a continuum of dressed dimer + hole.

Many-body treatment

The previous model tries to approximate a many-body problem with a 2-body description, so we have to extend and generalize it. We can still follow [55] in carrying on the description of the scattering with the 2-body transition matrix $\hat{T}^{2B}$ and introduce the medium effects by inserting the in-medium propagator for the two atoms (for more details see [33]). If one considers only a single scattering channel for the two atoms the generalized $\mathcal{T}$ matrix

$$\mathcal{T}(\mathbf{P},\omega) = \frac{2\pi}{m_r}\left[\frac{1}{a} - \frac{2\pi\Pi(\mathbf{P},\omega)}{m_r}\right]^{-1} \tag{3.69}$$

is obtained, where, for two particles with total momentum $\mathbf{P}$ and energy ω, the propagator of free atoms in medium is defined as

$$\Pi(\mathbf{P},\omega) = \int \frac{d\mathbf{p}}{(2\pi)^3}\left[\frac{1-f_\uparrow(\mathbf{p})-f_\downarrow(\mathbf{P}+\mathbf{p})}{\omega+i0^+-E_{\uparrow\mathbf{p}}-E_{\downarrow\mathbf{P}+\mathbf{p}}} + \frac{2m_r}{\mathbf{p}^2}\right] \tag{3.70}$$

$f_\sigma = [\exp(\beta E_{\sigma\mathbf{p}})+1]^{-1}$, $E_{\sigma\mathbf{p}} = \mathbf{p}^2/2m_\mathrm{r} - \mu_\sigma$ is the kinetic energy scaled by chemical potential μ_σ, and m_r is the reduced mass. The effects of the medium are thus included simply by adding the Fermi functions for the distribution of the fermions, avoiding in this way the scattering in

already occupied states. This minimal change has however great physical effects and allows one to obtain results comparable with what can be obtained in experiments and quantum Monte Carlo simulations.
The self-energy $\Sigma(\mathbf{p},\omega)$ describes the energy shift of minority particle due to interaction with the medium. We can expand $\Sigma(\mathbf{p},\omega)$ with respect to the number of holes created over the Fermi surface

$$\Sigma(\mathbf{p},\omega) = \Sigma^{(1)}(\mathbf{p},\omega) + \Sigma^{(2)}(\mathbf{p},\omega) + ... \tag{3.71}$$

where $\Sigma^{(n)}$ is a n-holes process. Taking only the leading term of this series corresponds to the so-called 1PH approximation. In such a way one has

$$\Sigma(\mathbf{p},\omega) = \Sigma^{(1)}(\mathbf{p},\omega) = \int \frac{d\mathbf{q}}{(2\pi)^3} f_\uparrow(\mathbf{q})\mathcal{T}(\mathbf{q}+\mathbf{p},\omega+E_{\mathbf{q}\uparrow}) \tag{3.72}$$

Then the Green's function for minority particle is then given by

$$G_\downarrow(\mathbf{p},\omega) = \frac{1}{\omega - E_{\mathbf{p}\downarrow} - \Sigma(\mathbf{p},\omega) + i0^+} \tag{3.73}$$

The polaron repulsive and attractive energies in figure 3.4 are computed for equal particles masses by Massignan et al. [55] as

$$E_\pm = \mathrm{Re}[\Sigma(0,E_\pm)] \tag{3.74}$$

- E_+ energy stands for "repulsive polaron": a quasiparticle formed, for positive scattering length, by the minority atom that repels the majority atoms that surround it. In the BEC limit $E_+ = 2\pi a n_\uparrow/m_r$, it increases moving toward Feshbach resonance, and reaches the order of ϵ_F. Approaching the resonance, however, this state becomes unstable to decay and thus ill-defined.

- E_- energy stands for "attractive polaron": a quasiparticle formed by the minority atom that attracts the majority atoms that surround it. In the BCS limit $E_- = 2\pi a n_\uparrow/m_r$, it becomes more negative by increasing $1/k_F a$. For equal masses and for a broad resonance ($k_F R^* \ll 1$), $E_- \sim -0.6E_F$ at unitarity ($1/k_F a = 0$).

The quasiparticle residue and effective mass of the polaron can be obtained from eq.(3.72) as

$$Z_\pm = [1 - \partial_\omega \Sigma(0,\omega)|_{\omega=E_\pm}]^{-1} \tag{3.75}$$

$$m^*_\pm = \frac{m}{Z_\pm}\left[1 + \left(\frac{m}{p}\frac{\partial \mathrm{Re}\left[\Sigma(\mathbf{p},E_\pm)\right]}{\partial p}\right)\Big|_{p\to 0}\right]^{-1} \tag{3.76}$$

In the proximity of a quasi-particle pole of quasiparticle, the Green's function can be approximated as

$$G_\downarrow(\mathbf{p},\omega) \sim \frac{Z_\pm}{\omega - E_\pm - p^2/2m^*_\pm + \mu_\downarrow + i0^+} \tag{3.77}$$

so this quasiparticle can be seen as a free particle with renormalized mass and spectral weight. Moreover, as shown in grey in figure 3.4, a continuum of states between the two polaron branches appears in the BEC limit. It is a dressed dimer + hole continuum, where the dressed dimer is made by one minority plus one majority atom (extracted from the Fermi sphere) leading to a continuum of excitation of spectral width $\sim E_F$.

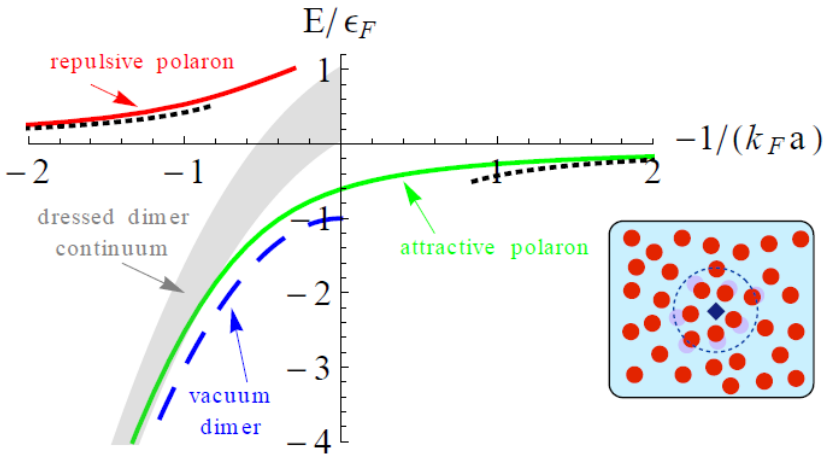


Figure 3.4: Plot of energy spectrum of single impurity in a Fermi sea obtained from the 1PH approximation. We are interested in unitarity where the attractive polaron is the physical state. Mean-field results are also shown for comparison (dotted black lines). The dashed line is the dimer energy in the absence of the Fermi sea. Figure from [55] .

Variational ansatz

The 1PH approximation corresponds to a variational ansatz [56], expanding a many-body state over the number of particle-hole excitations in Fermi sea. The lowest order will be (1PH) for the many-body state is

$$|\psi\rangle = \sqrt{Z}a^\dagger_{0\downarrow}|FS_N\rangle + \sum_{q<k_F<k} \phi_{\mathbf{q},\mathbf{k}} a^\dagger_{\mathbf{q}-\mathbf{k}\downarrow} a^\dagger_{\mathbf{k}\uparrow} a_{\mathbf{q}\uparrow}|FS_N\rangle \tag{3.78}$$

where $a_{\mathbf{k}\sigma}$ annihilates a fermion of momentum $\mathbf{k}$ and spin σ, $|FS_N\rangle$ is the Fermi sea of $N_\uparrow$ fermions, and $\{Z, \phi_{\mathbf{q},\mathbf{k}}\}$ are variational parameters. In particular Z is the quasiparticle residue. It is possible to write down a variational ansatz also for the dressed dimer. Expanding the many-body wavefunction with respect to particle-hole excitation number in Fermi sea, one writes

$$|\Psi\rangle = \sum_{k>k_F} \phi_{\mathbf{k}} a^\dagger_{-\mathbf{k}\downarrow} a^\dagger_{\mathbf{k}\uparrow}|FS_{N-1}\rangle + \sum_{q<k_F<k,k'} \phi_{\mathbf{q},\mathbf{k},\mathbf{k}'} a^\dagger_{\mathbf{q}-\mathbf{k}-\mathbf{k}'\downarrow} a^\dagger_{\mathbf{k}\uparrow} a^\dagger_{\mathbf{k}'\uparrow} a_{\mathbf{q}\uparrow}|FS_{N-1}\rangle \tag{3.79}$$

where $|FS_{N-1}\rangle$ is the $N_\uparrow - 1$ fermions Fermi sea, and $\{\phi_{\mathbf{k}}, \phi_{\mathbf{q},\mathbf{k},\mathbf{k}'}\}$ are variational parameters. This system of one spin-$\downarrow$ in a spin-$\uparrow$ gas, has a polaron ground state in the BCS limit, and a dressed dimer in a N-1-spin-$\uparrow$ Fermi sea in the BEC limit. When the ground state is a polaron, a zero-momentum continuum of states is also present. It corresponds to a finite momentum polaron, with an opposite momentum particle-hole excitation on the Fermi sea [55].

3.4.2 Beyond the Fermi polaron

Until now, we have described the theory suitable for very high polarization. In particular, if only spin-$\uparrow$ are present (fully polarized system), the system will behave as an ideal Fermi gas. When we add also spin-$\downarrow$ fermions, then we have to analyse the system with Fermi liquid theory in general. For a single spin down, but also for few of them, we turn to the Fermi polaron theory from the previous subsection.

If now we switch to zero polarization ($p = 0$), corresponding to a perfect matching of populations, we expect that pairs of fermions are formed, and, since we are assuming zero temperature, they will condense in a homogenous superfluid phase. In this phase, BCS theory holds. If we now change a little the population, we expect that the population imbalance perturbs this state, decreasing pairing and then also superfluidity. Even if the superfluid state can sustain a small imbalance, for p sufficiently large, BCS theory breaks down. Then if we try to connect the two regions of the phase diagram, starting from high polarization and decreasing p, there must be some point, in the polarization axis, where the phase changes, with the system presenting a phase transition from normal to superfluid. As discussed in [1] and [57], depending on the interaction strength this transition could be first or second order.
In the case of a second order transition , critical fluctuations in the proximity of the critical point may lead to strong interaction effects in Fermi liquid phase. In particular, in [1] Pini et al. have found that by progressively decreasing the polarization from the polaronic limit, the effective masses of spin down and up quasi-particles increase, until they diverge at a critical polarization at which the stable phase of the system becomes a Fulde-Ferrel-Larkin-Ovchinnikov (FFLO) superfluid. Such a phase can sustain a finite polarization because it allows the Fermi surfaces to be partially matched, and pairs condensate with a finite value of centre-of-mass momentum. In next subsection we will briefly summarize the results of [1] on the evolution of the polarized unitary Fermi gas from the polaronic limit down to the quantum critical point QCP found in FFLO superfluid phase.

Fully self-consistent T-matrix results

In [1] the previous approach based on the many-body extension of the two-body T-matrix is made fully self-consistent. In practice, the ladder series of Feynmann diagrams for the T-matrix is constructed in terms of self-consistent Green's functions, which are dressed by a T-matrix self-energy. The theory is formulated for convenience on the imaginary frequency axis, where the previous expression for quasi-particles residue Z_σ and effective mass m^*_σ read:

$$Z_\sigma = \left[1 - \frac{\partial \mathrm{Im}\Sigma_\sigma(k_{F\sigma}, i\omega)}{\partial\omega}\bigg|_{\omega=0^+}\right]^{-1} \tag{3.80}$$

$$\frac{m}{m^*_\sigma} = Z_\sigma \left[1 + \frac{m}{k_{F\sigma}} \frac{\partial \mathrm{Re}\Sigma(\mathbf{k}, i0^+)}{\partial|\mathbf{k}|}\bigg|_{|\mathbf{k}|=k_{F\sigma}}\right] \tag{3.81}$$

and the T-matrix self-energy, is given by

$$\Sigma_\sigma(k) = -\int \frac{d\mathbf{Q}}{(2\pi)^3} \int \frac{d\Omega}{2\pi} \Gamma(Q) G_{\bar\sigma}(Q - k) \tag{3.82}$$

where $Q = (\mathbf{Q}, i\Omega)$, $G_\sigma(Q-k) = [G_{0\sigma}(k)^{-1} - \Sigma_\sigma(k)]^{-1}$ is the dressed Green's function ($\bar\sigma$ being the opposite of σ), and $\Gamma(Q)$ is the particle-particle propagator

$$\Gamma(Q) = -\left[\frac{m}{4\pi a} + R_{\mathrm{pp}}(Q)\right]^{-1} . \tag{3.83}$$

where R_{pp} is the renormalized particle-particle bubble

$$R_{\mathrm{pp}} = \int \frac{d\mathbf{k}}{(2\pi)^3} \left[\int \frac{d\omega}{2\pi} G_\sigma(k) G_{\bar\sigma}(Q - k) - \frac{m}{k^2}\right] \tag{3.84}$$

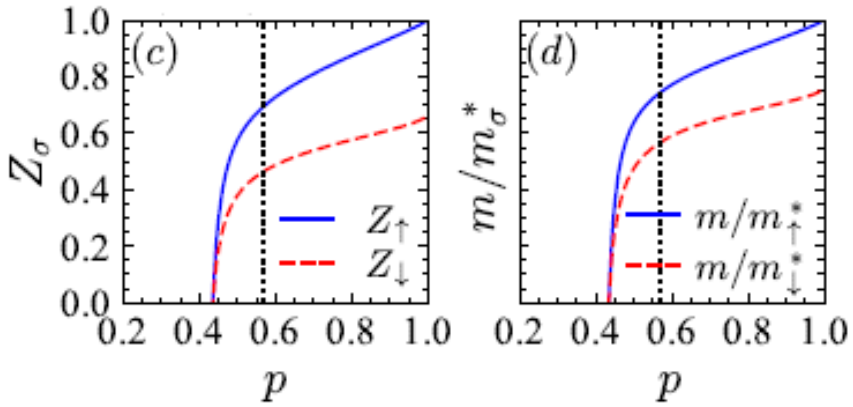


Figure 3.5: Quasiparticle residue Z_σ and effective mass m^*_σ for the two spin components as function of polarization p for a unitary Fermi gas. Figure from [1].

One can see the behaviour of the quasiparticle residues and effective masses as functions of polarization in Fig.3.5, where we restrict our attention to unitarity. For $p \to 1$ the polaronic limit holds, the majority particles are essentially not interacting and the minority particle is dressed by interactions to form an attractive polaron. This situation changes continuously; the quasi-particle residues and the inverse effective masses decrease linearly with polarization until a crossover value p^* is reached. After this value the system enters a critical region, where the dependence with polarization is much stronger, until at a critical value p_c, Z_σ vanishes and m^*_σ diverges. The dotted line in Fig.3.5 corresponds to the value of p^* determined as a 5% deviation from the linear behaviour. So around the quantum critical point there is a breakdown of Fermi liquid theory, evidenced by the vanishing of the quasiparticle residues and the divergence of the effective masses for both spin species.
It is important to inspect also the behaviour of the pair propagator, that in this approach corresponds to the pairing susceptibility. Pini et al. monitored the momentum dependence of such quantity, looking for its divergence

$$\left[\Gamma\left(|\mathbf{Q}| = Q_0, i\Omega = 0\right)|_{p=p_c}\right]^{-1} = 0 \tag{3.85}$$

When the pairing susceptibility diverges at a finite pair wave vector Q_0, there is a second order phase transition towards a FFLO phase; otherwise, for $Q_0 = 0$, the transition will be toward a polarized superfluid. The results for $\Gamma^{-1}(Q;0)$ at p_c are shown in fig.3.6a, while in fig.3.6b one can see the value of the momentum Q_0 at the divergence of the pairing susceptibility.
Note that increasing the critical polarization corresponds to increasing the coupling strength $(k_F a)^{-1}$. One sees from Fig.3.6a that for increasing p_c (i.e increasing coupling) the minimum of $\Gamma^{-1}(Q;0)$ changes from a minimum at finite Q_0 to a minimum at zero momentum with a sudden jump; evidencing a sudden change from a regime where the transition is to FFLO to a regiem where the transition is to a poalrized "BCS" superfluid.

Zero temperature phase diagram

The resulting phase diagram obtained from eq.(3.85) is presented in figure 3.7a with a comparison with experimental and QMC data. It covers a range of interaction strengths, with unitarity $[(k_F a)^{-1} = 0]$ just in the middle of this range. On the left of the lifshitz point (L), the transition is toward a polarized BCS superfluid (pBCS) namely with pairing at zero momentum. Only

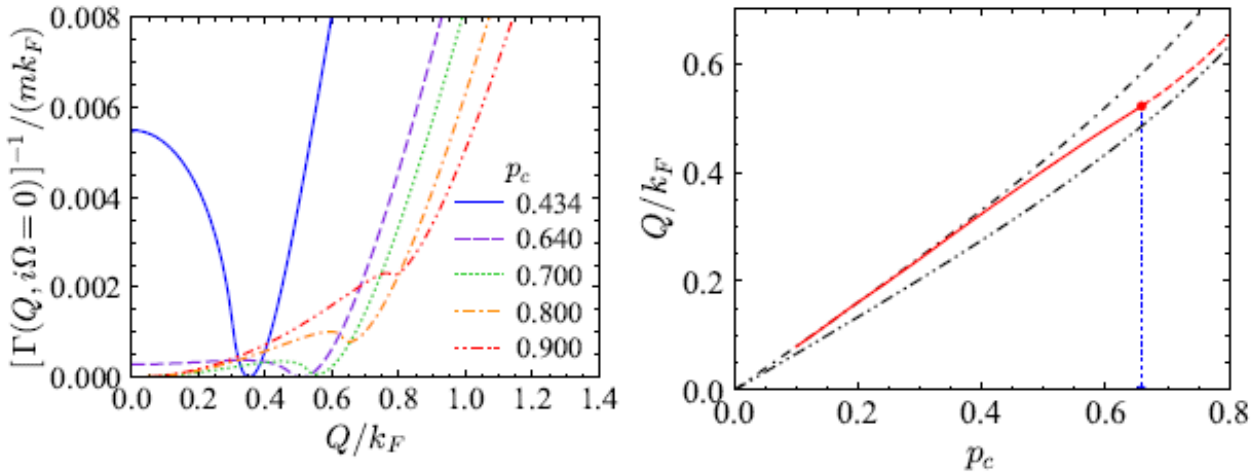


Figure 3.6: a) Momentum dependence of the inverse of the pair-propagator at zero frequency at the critical polarization p_c, for different values of p_c (see legend), corresponding to $(k_F a_F)^{-1}$ = 0.00; 0.59; 0.69; 0.83; 0.99, from top to bottom. b) Pair wave vector Q_0 (in units of k_F) vs polarization p_c at the superfluid phase transition. The dashed line for $p_c > p_L$ corresponds to the secondary minimum of $\Gamma(Q;0)^{-1}$ (the absolute minimum being at $Q_0 = 0$). The dash-dotted line represents the Fulde-Ferrell mean-field result $Q_0 = 1.2|k_{F\uparrow} - k_{F\downarrow}|$ [60], while the dashed-double-dotted line corresponds to $Q = |k_{F\uparrow} - k_{F\downarrow}|$. Figures from [1].

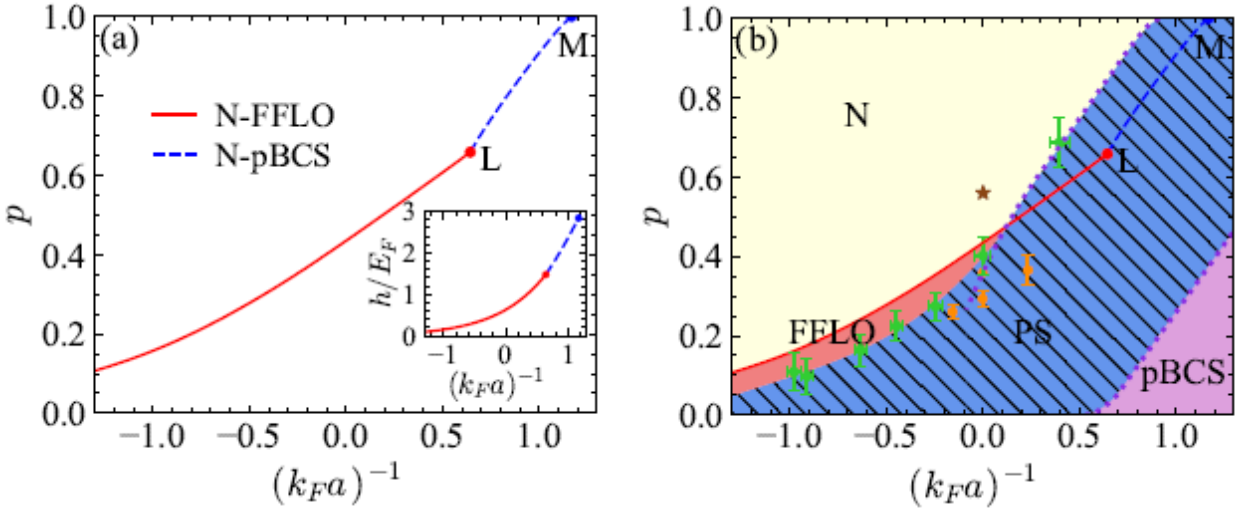


Figure 3.7: Polarization-vs-coupling phase diagram at zero temperature from [1]. (a) Phase diagram calculated within the self-consistent t-matrix approach for the second order phase transition between the normal (N) phase and a FFLO superfluid (full line) or a polarized BCS (pBCS) superfluid (dashed line). At the Lifshitz point (L) the instability changes from FFLO to polarized BCS superfluid. Inset: corresponding critical $h = (\mu_\uparrow - \mu_\downarrow)/2$ vs coupling. (b) Phase diagram taking into account phase separation (hatched region) between the normal and pBCS superfluid phases as obtained from the experimental works [58][59](circles) and [60] (diamonds), as well as QMC calculations [61] (dotted lines). Star: result for the N-FFLO phase transition at unitarity from [62]. In both panels the point M corresponds to the polaron-to-molecule transition.

when the transition is to the FFLO phase one observes a complete breakdown of Fermi liquid properties with $Z_\sigma \to 0$ and $m^*_\sigma \to \infty$ at the QCP, while when the transition is toward pBCS the quasiparticle residue Z_σ remains finite.
In fig.3.7b the second order transition line of fig.3.7a is compared with an interpolation of of experimental and QMC data accounting for phase separation. One can see that only a small region of the phase diagram is available to the FFLO state, pointing to the difficulty for an experimental evidence of such a phase in ultracold gas experiments.
Still according to Pini et al., at unitarity the QCP to the FFLO phase could be reached before phase-separation sets in, by lowering progressively the polarization from the polaronic limit. We wish to investigate this scenario with Quantum Monte Carlo simulations. It will be interested in particular in determining the effective masses in the normal phase of the polarized Fermi gas, and search for indications of a possible divergence at a critical polarization.

CHAPTER 4

MONTE CARLO METHOD

There are many different numerical methods to solve integrals but most of them are effective only when the dimension of the system is lower than four. When the problem requires high dimensional integration they lose their efficiency. Differently from this situation the Monte Carlo (MC) method has no dependence on dimension, the precision scales always as $N^{-1/2}$ (where N is the total number of sampled points). So it is really effective in computing observables in a high-dimension phase space.
This method was also introduced to solve the Schrödinger equation, first with variational Monte Carlo, by McMillan [63]. In this case, exploiting the variational principle, it is possible to obtain an integral to be solved with MC that is an upper bound of the ground state energy. Then it was developed also in other forms, by means of Green's functions to solve the problem as a diffusion equation, or by exploiting the path integral formulation of quantum mechanics or diagrammatic methods.
Often it turns out that this approach is the only accessible tool for studying sophisticated problems. Usually an analytically solvable model costs some assumptions and approximations, that are sometimes quite severe, which can be relaxed in Quantum Monte Carlo (QMC). Often an attempt can be made to rely on small parameters of the system to construct a perturbative theory of a phenomenon, and also in these cases QMC can avoid such restrictions. When we are interested in exploring the quantum properties of systems, QMC methods can solve the Schrödinger equation for ground state or excited ones. Since the quantum effects manifest at most at the lowest temperatures, we will consider the system in its ground state. Thus we choose the formulation that exploits Green's functions, which is called Diffusion Monte Carlo or DMC, and consists in letting the system evolve by transforming the initial state with successive iterations of the propagator (Green's function) approximated for very short imaginary-time intervals. The repetition for many times allows to arrive at a final state, that differs from the first for a finite imaginary-time. If this interval is long enough, it is possible to project the system to its ground state.
This method could be exact, up to statistical error, in calculating ground state energy for a bosonic system, since a bosonic ground-state wavefunction can generically be interpreted as a probability density, being semidefinite positive. Conversely in order to study a fermionic system one has to introduce a Fixed-Note Monte Carlo technique, that allows to tackle the sign changes of the wavefunction, intrinsic of fermionic systems, by assuming a reference "nodal surface". In general this approach gives an upper bound to the ground state energy, but with a good choice of the trial wave function the difference can be significantly minimized.

4.1 Brief history of Monte Carlo method

Let us start considering a brief analysis of the history of Monte Carlo [64][65]. The name of this method was coined by Nicolas Metropolis in 1949 [66], before this the older name was "statistical sampling". It was born as a method to estimate integrals which could not be performed by other means. Integrals over poorly-behaved functions and integrals in high-dimensional spaces are two areas in which the method has traditionally proved profitable, and indeed it is still an important technique for problems of these types.
A famous example of the use of Monte Carlo is the experiment known as Buffon's needle [67], in which the constant π is determined by dropping a needle onto a surface with evenly spaced lines. It's only thanks to the advent of mechanical calculating machines at the end of XIX century that numerical methods took a large step forward. They enabled the statistical sampling techniques. An early example of what was effectively a MC calculation of the motion and collision of the molecules in a gas was described by William Thomson (later Lord Kelvin). Then numerical methods had a great importance during World War.
The first real applications of statistical sampling method in physics seem to have been those of Enrico Fermi, who was working on neutron diffusion in the early 1930s. The important development of the method took place at Los Alamos National Laboratory in New Mexico by Nick Metropolis, Stainslaw Ulam and John von Neumann. It seems that Stan Ulam reinvented Fermi's statistical sampling method after the idea of finding numerically the possibility to win a card game: he abandoned the aim of analytical solution by means of combinatorial calculus. A machine could play the game much faster than him and count the victories. The same concept is valid for physical phenomena. When the modifications suggested by Richard Clippinger and realised by Klara Dan von Neumann were ready it was the end of 1947, then finally Metropolis and von Neumann set to work on the Monte Carlo calculation. This first calculation was followed by other works on hard-sphere gasses, and the article published in 1949 claimed the world's attention.
The publication in 1953 [68] by Nick Metropolis, Marshall and Arianna Rosenbluth, and Edward and Mici Teller, described for the first time the MC technique that has come to be known as Metropolis algorithm. It was the first application of the thermal "importance sampling" technique. It has been frequently utilised since then, first in conjunction with the variational principle for both classical and quantum issues, first for bosons [63] and then for fermions [69]. It was later improved to solve complex systems more accurately.

4.2 Basic aspects of the method

4.2.1 Definition of Monte Carlo method

Monte Carlo (MC) can be intended as a general strategy, more then a technique. It can be applied to a large variety of different problems, where analytic or deterministic solution is out of question. Kalos [65] defines it as:
One method that involves deliberate use of random numbers in a calculation that has the structure of a stochastic process. By stochastic process we mean a sequence of states whose evolution is determined by random events. In a computer these are generated by random numbers.
He assumes a difference between a generic simulation and Monte Carlo. While, in first case, one tries to transcript into computing terms a natural stochastic process, Monte Carlo is the solution by probabilistic method (deliberate use) of non probabilistic problems. In general these two things are weakly distinguishable. In fact if one considers the emission of radiation

from atoms, and its interaction with matter, one has a natural stochastic process. However one can describe the process in average by writing mathematical equations, which are solvable by random sampling with Monte Carlo method.
Theoretical and technological advances lead to upgrades of the method, both in speed and effectiveness, turning it into an efficient numerical method for a wide class of problems. So the method confirms the expectations of Metropolis and Ulam. They were the first to call it Monte Carlo [66], defining it as:
The Monte Carlo method is an iterative stochastic procedure, consistent with a defining relation for some function, which allows an estimate of the function without completely determining it.
Let us analyse this sentence following the work by Gubernatis, Kawashima, & Werner [70]. The method is an *iterative stochastic procedure*, meaning the procedure is applied over and over. Then one can obtain a large number of measurements, connected as a chain of statistical events. This leads to a convergence to the best estimate in statistical sense by means of law of large numbers. It is a stochastic procedure, so random number are generated according to one *defining relation*. That rule corresponds to differential and integro-differential equations, *for some function* which is the solution to these equations.

4.2.2 Random events and basic statistics

Let us just assume the existence of random events, and that we are able to organize a computer program to produce effective equivalents of natural random events. In general they can be random because of some composition laws. It is difficult to distinguish which natural random events are elementary, from composite ones. But since the distinction depends on one's state of knowledge and the depth of the analysis given to the problem, we can identify each of our fundamental events as elementary. We will associate to each of them a countable set of random outcomes, E_i, with relative probability, p_i.
Even if sometimes it could be not possible, we will assume the existence of a map that connects each outcome to a numerical value. This connection could be also with a boolean number, that leads to a specific branch in a program. As explained in more detail by Kalos [65], but also in many other textbooks, we call random variable a real number x_i associated to each outcome E_i. The expectation of this random variable x (the stochastic mean value) is defined as

$$E(x) = \sum_i p_i x_i = \langle x \rangle \tag{4.1}$$

If we consider now some real-valued function $g(x_i) = g_i$, if x_i is a random variable, then also $g(x_i)$ is a random variable, and one can define its expectation in a similar way

$$E(g(x)) = \sum_i p_i g(x_i) = \langle g(x) \rangle \tag{4.2}$$

One can have an idea of the dispersion of the random variable by computing the square root of the variance, that is the second central moment. For a random variable with average μ it corresponds to

$$\mathrm{VAR}[x] = \langle (x-\mu)^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle = \sum_i p_i (x_i - \mu)^2 = \sum_i p_i x_i^2 - \langle x \rangle^2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{4.3}$$

similar is the variance for a function $g(x)$.

Let us now move to consider a set of random variables $x_1, ..., x_n, ...$ all drawn at random from the same probability density function $f(x)$. A function G may be defined

$$G = \sum_{n=1}^{N} \lambda_n g_n(x_n) \tag{4.4}$$

Each g_n is a random variable then it is also their sum. Since the expectation value is linear, in this case we have

$$E[G] = \langle G \rangle = \sum_{n=1}^{N} \lambda_n \langle g_n(x_n) \rangle \tag{4.5}$$

and, if all x_n are independent, we also have

$$\mathrm{VAR}[G] = \sum_{n=1}^{N} \lambda_n^2 \mathrm{VAR}[g_n(x_n)] \tag{4.6}$$

Since G is a function of the observation $g(x)$, we can say it is an estimator [71].
Let us say something more specific about estimators. If we take a function of some quantity θ, it can be a satisfactory approximation of an observable O if it is not expected to fluctuate far from O

$$\langle (\theta - O)^2 \rangle / O^2 \ll 1 \tag{4.7}$$

Where the numerator can be written as

$$\langle (\theta - O)^2 \rangle = \langle (\theta - \langle \theta \rangle)^2 \rangle + \langle (\langle \theta \rangle - O)^2 \rangle + 2 \langle (\theta - \langle \theta \rangle) \rangle \langle (\langle \theta \rangle - O) \rangle = \langle (\theta - \langle \theta \rangle) \rangle + (\langle \theta \rangle - O)^2 \tag{4.8}$$

The quality of θ depends on the variance of θ and on the departure of its mean from O. The quantity $\langle \theta \rangle - O$ is called the bias of the estimator. An unbiased estimator has the correct mean value for every number N of samples. When an unbiased estimator is not possible, it is important that its bias decreases to zero with increasing N.
An estimator θ is consistent with the quantity O if it converges with probability 1 as N approaches infinity, to O

$$P\left\{ \lim_{N \to \infty} \theta(x_1, ..., x_N) = O \right\} = 1 \tag{4.9}$$

In particular $\langle G \rangle$ can be used to obtain an approximation of the expectation $\langle g(x) \rangle$. Now if all coefficients λ_n are identical and equal to $1/N$, and also all function g_n are the same, the expectation becomes

$$E[G] = \langle G \rangle = \langle g(x) \rangle \tag{4.10}$$

The function G is the arithmetic average of g, they have the same mean. The variance is

$$VAR[G] = VAR[\sum_{n=1}^{N} g_n(x_n)/N] = \sum VAR[g(x)]/N^2 = VAR[g(x)]/N \tag{4.11}$$

So as the number of samples of x increases, the variance of the mean decreases as $1/N$. This leads to the central idea of Monte Carlo evaluation of integrals: use a sum to approximate an integral like

$$\langle g(x) \rangle = \int_{-\infty}^{\infty} g(x) f(x) dx = E\left[\frac{1}{N} \sum_{n=1}^{N} g_n(x_n) \right] \tag{4.12}$$

It is a series of random variables from a distribution f. One then evaluates every time $g(x)$, and finally one gets the estimate of the integral from the arithmetic mean of all the values of g. The variance of this estimate decreases as the number of terms increases.

Law of large numbers and Chebychev inequality

Let us try now to discuss more generally the previous result that will be called the law of large numbers of probability theory [71]. For simplicity we will assume that the mean and the variance of our random variable always exist. Let us suppose to have a set of independent and identically distributed random variables $x_1, ..., x_N$, so that the expectation of each x_i is the same μ. Then their average value is

$$\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4.13}$$

that for $N \to \infty$ converges in probability [72] to μ. However one can see, by means of Bienaymé-Tchebycheff inequality [71], that large deviations from the expectation value decrease in probability when the number of points increases, in fact for our random variable G

$$P\left[|G - \langle G\rangle| \geq (k\mathrm{VAR}[G])^{1/2}\right] \leq \frac{1}{k} \tag{4.14}$$

where k is a positive number. Since we can make N as big as we want, the variance of G becomes as small as we want, so our estimate get a very small probability of a large deviation relative the real value. Otherwise, if we increase the sample size, we will have the same probability to obtain values of G in smaller regions near $\langle g\rangle$. By these laws, the sample mean $\bar{x}_N$ is a consistent and unbiased estimator of the mean μ.

Central limit theorem (CLT)

A much stronger statement than the Bienaymé-Tchebycheff inequality about fluctuations of G is given by the central limit theorem of probability (CLT) [71]. For a fixed value of N one can find a pdf that describes the values of G that will occur. But, as $N \to \infty$,this theorem shows that there is a specific limit distribution: the normal distribution. Set

$$\begin{aligned} G_N &= \frac{1}{N}\sum_{i=1}^{N} g(x_n) \\ \mathrm{VAR}[G_N] &= \mathrm{VAR}[g]N \end{aligned} \tag{4.15}$$

and

$$t_N = (G_N - \langle g\rangle)/\left[\mathrm{VAR}[G_N]\right]^{1/2} = \frac{\sqrt{N}(G_N - \langle g\rangle)}{\left[\mathrm{VAR}[g]\right]^{1/2}} \tag{4.16}$$

then

$$\lim_{N\to\infty} P\{a \leq t_N \leq b\} = \int_a^b \frac{\exp[-t_N^2/2]}{\sqrt{2\pi}} \tag{4.17}$$

Letting $\sigma^2 = \mathrm{VAR}[g]$, the previous element inside the integral can be written as

$$f(G_N) = \frac{1}{\sqrt{2\pi(\sigma^2/N)}}\exp\left[-\frac{N(G_N - \langle g\rangle)^2}{2\sigma^2}\right] \tag{4.18}$$

As $N \to \infty$ the observed G_N assumes a Gaussian shape around $\langle g\rangle$ and one can predict the probability of deviations measured in units of σ. So we have a specific distribution to describe

G_N, even if this is correct only asymptotically. The variance used above may be estimated from the observed independent values of $g(x_n)$. Let us apply the definition

$$\left\langle \frac{1}{N}\sum_n g^2(x_n) - \left[\frac{1}{N}\sum_n g(x_n)\right]^2 \right\rangle = \langle g^2\rangle - \frac{1}{N^2}\left\langle \sum_n g^2(x_n) + \sum_{n,m\neq n} g_n g_m \right\rangle \tag{4.19}$$

Using the independence of g_n and g_m in evaluating $\langle g_n g_m\rangle$, we find the right-hand side equal to

$$\left(1-\frac{1}{N}\right)\langle g^2\rangle - \frac{N(N-1)}{N^2}\langle g\rangle^2 = \frac{N-1}{N}\mathrm{VAR}[g] \tag{4.20}$$

Thus an estimator for σ^2 is

$$\sigma^2 \simeq \frac{N}{N-1}\left\{\frac{1}{N}\sum_n g^2(x_n) - \left(\frac{1}{N}\sum_n g(x_n)\right)^2\right\} \tag{4.21}$$

We assumed before the possibility to have at our disposal a set of truly random variables. In a real calculation they are not random but pseudorandom, generated through a generation routine. We choose to use that method for the possibility to repeat the particular run of the program, and so we can be able to debug a computer code. Real random numbers lead to very low probability of identical calculation, and the recurrence of an error would be left to chance. Also it is important to have the possibility to repeat the same calculation after some changes of the program.

4.2.3 Definition of sampling

Let us discuss now one more basic concept in Monte Carlo: the random sampling. It bridges probability theory and statistics. We are assuming the existence of some experiment that produces a realization of a random variable X from a set of possible outcomes. Consider a value x of variable X after experiment, which lays in probability space Ω_0, $X \in \Omega_0$, and it follows a probability density function $f(X)$, where

$$\int_{\Omega_0} f(X)dX = 1 \tag{4.22}$$

If N experiments are performed simultaneously (an ensemble), associated with the i-th experiment is the random variable X_i. They are all identical to the random variable X and by assumption have the same distribution as X; that is, $f_i(X_i) = f(X)$. The joint distribution of these N random variables is $f(X_1, X_2, \ldots, X_N)$. They are statistically independent if

$$f(X_1, X_2, \ldots, X_N) = f_1(X_1)f_2(X_2)\ldots f_N(X_N) \tag{4.23}$$

The ensemble of N experiments thus produces a set of outcomes $\chi = \{\chi_1, \chi_2, ..., \chi_N\}$ such that the i-th random variable takes the value $X_i(\chi) = X(\chi_i) = x_i$. A sampling procedure is an algorithm that can produce a sequence of values of X, $x = \{x_1, x_2, \ldots\}$ such that for an $\Omega \subseteq \Omega_0$

$$P\{x_k \in \Omega\} = \int_\Omega f(x)dx \leq 1 \tag{4.24}$$

Both independence and identical distributions are necessary for random sampling. In this thesis it will be necessary to sample a random variable from a gaussian distribution. This can be done

exploiting the Box-Muller algorithm [73]. It states the possibility to sample two independent gaussian random variable together

$$f(y_1, y_2) = \rho(y_1|0,1)\rho(y_2|0,1) = \frac{1}{2\pi}\exp[-(y_1^2 + y_2^2)/2] \tag{4.25}$$

from two independent uniformly distributed random variables ξ_1 and ξ_2 by transformation in polar coordinates. So one has to consider

$$\begin{aligned} y_1 &= [-2\log\xi_1]^{1/2}\cos(2\pi\xi_2) \\ y_2 &= [-2\log\xi_1]^{1/2}\sin(2\pi\xi_2) \end{aligned} \tag{4.26}$$

Then the equation can be linearly transformed to any μ and any σ.

4.2.4 Statistical physics

When one is dealing with condensed matter systems, usually there are too many elementary constituents to solve analytically the equations of motion. Even if they obey quite simple equations, which can be solved easily for non interacting constituents. The problem arises when these constituents are left free to interact. This often leads to an impossible problem to solve analytically. However, independently with respect to the number of equation there are, the system's average behaviour is predictable. Solving the problem by approaching the system in a probabilistic manner is the aim of statistical mechanics. In that perspective, the likelihood that the system will be in one condition or another is more important than the precise result. The majority of the potential states for big systems are typically highly improbable, giving us a great degree of confidence in the behaviour that the actual system would exhibit.
The systems that will be studied in this thesis are quantum systems specified by a Hamiltonian operator acting on Hilbert space. We will show later, however, that quantum Monte carlo (QMC) methods typically map a quantum system to a classical one with a suitable projection to a configuration basis, so that considerations valid for classical Monte Carlo methods are also applicable to QMC methods. For hamiltonian systems energy is conserved, so the system will stay forever in the same state or at last it can switch to degenerate states. However in classical thermal systems there is another component: the thermal reservoir. An external system that constantly exchange energy with our hamiltonian system toward equilibrium. This reservoir is a weak perturbation, negligible in calculating the energy levels of the system but effective in changing its state. By following [64], we can analyse how the effects of reservoir can be incorporated through dynamics, a law for changing state with a nature imposed by the particular perturbation.
Suppose now that the system moves from one initial state μ to a final state ν with a rate $R(\mu \to \nu)$, so the probability of change in time dt is given by $R(\mu \to \nu)dt$. A transition rate can be defined for every possible final state. So after a small time interval, the system can move to one of a large possibility of states. Now let us introduce a weight for the system to be in an initial state μ, at time t, as $w_\mu(t)$. The evolution of the weight in time is defined by a master equation (a set of equations, one for each state)

$$\frac{dw_\mu}{dt} = \sum_\nu [w_\nu(t)R(\nu \to \mu) - w_\mu(t)R(\mu \to \nu)] \tag{4.27}$$

It depends on the rate of transition in both directions. The probabilities appearing in previous equation must satisfy the condition

$$\sum_\mu w_\mu(t) = 1 \tag{4.28}$$

for all t, that means that the system must always be in one state.
According to statistical mechanics and the definitions from the preceding subsection, when one is interested in computing the best estimate of an observable O, each possible state of the system O_μ must be considered

$$\langle O \rangle = \sum_\mu O_\mu w_\mu(t) \tag{4.29}$$

How this expression is related to experimental measurement of O can be explained in two ways. First one can imagine to have a large number of identical and independent copies, each with its own reservoir. Then performing the average from an instantaneous measure of O from all copies (the ensemble) one gets a good estimate of $\langle O \rangle$. When one has at disposal only one copy of the system he has to consider a time average. This quantity will be similar to $\langle O \rangle$ if in a certain interval of time the system passes through a representative selection of the states in the probability distribution w_μ. Longer time interval will lead to a better estimate of the average. The problem for the second approach is how to be sure to catch a representative selection of states. Moreover, for non-equilibrium systems, there is also the possibility that the probability distribution w_μ changes during measurement.
Let us consider simply an equilibrium state, since the main goal of the thesis is to simulate a fermionic system in equilibrium with Monte Carlo techniques. In that condition the rates of change dw_μ/dt will all vanish, and so all weights will take constant values for the rest of time. Since the master equation is first order with real parameters, and since the variables w_μ are constrained to lie between 0 and 1, that prohibits exponentially growing solutions, all systems governed by this equation must came to equilibrium in the end.
The important point is that one has to reach the equilibrium probabilities p_μ, from which the observable can be averaged as

$$\langle O \rangle = \sum_\mu O_\mu p_\mu \tag{4.30}$$

Let us go on with the discussion of a solution in a probabilistic fashion of a problem. We have also to deal with fluctuations in observable quantities. Typically after the computation of the average one can compute also the mean square deviation as

$$\langle (O - \langle O \rangle)^2 \rangle = \langle O^2 \rangle - \langle O \rangle^2 \tag{4.31}$$

the RMS fluctuation of this quantity is just the square root of O. It scales like the square root of the size of the system, so the relative fluctuation of an extensive quantity decreases as the size to minus one half. Then in the limit of a very large system, the so-called thermodynamic limit, fluctuations become negligible. This is the perfect regime to analyse condensed matter systems. Unfortunately in computer simulation this limit is inaccessible, and the best one can do is making an efficient algorithm to simulate the largest possible system in the available computer time, hoping to approximate sufficiently well the thermodynamic limit.
Statistical mechanics describes the system starting from the partition function Z but it assumes the possibility to perform infinite sums and integrals of infinite dimension. Typically computational methods are based on putting the system on a lattice of finite size, in order to reduce infinite sums and integrals to finite operations. This procedure generates an error that must be adjusted to have a correct comparison with thermodynamic limit.
In a classical Monte Carlo calculation one tries to reproduce the value of an observable during experiment and compute the time average, making the model passing through a variety of states in such a way that the probability of being in a particular state at time t corresponds to the respective weight $w_\mu(t)$. One must devise a plan to determine the transitional set of rates

in order to do this. After that, one utilizes them to decide which states the system will visit during the simulation and computes the average from those. The benefit of this method is that it just requires a small fraction of possible states to estimate the observables. The drawback of this non-infinite number of points is the statistical error in calculation. In the chapters that follow, the analysis of correction due to time constraints will be covered.

4.3 Application of the method

Let us consider again first a classical system. In the thermodynamic limit and zero temperature typically one can determine all ground state properties if energy and state are known. When temperature is finite one can be interested in the equation of state, with emphasis on the behaviour of the system near phase transitions. To study this phenomena we need to extrapolate the system to the thermodynamic limit and look for singular behaviour in the free energy or in derived thermodynamic functions.
If one considers the Ising model in one dimension, it can be solved with an exact expression for the partition function. Transfer matrix technique [74] yields the exact solution of this and other models, for example, the zero-field two-dimensional Ising model. However with Monte Carlo methods, the most convenient approach to thermodynamics is the computation of averages from the Boltzmann probabilities for the possible configurations of the microscopic variables. So to estimate one or more observables O, through calculation of expectation value $\langle O \rangle$, the ideal route is that of averaging the quantity of interest, performed over all states of the system weighted with a probability

$$\langle O \rangle = \frac{\sum_\mu O_\mu w_\mu}{\sum_\mu w_\mu} \tag{4.32}$$

In our example if O represents a physical quantity of interest, that is a function of the Ising variables, we can compute the temperature dependent average (thermal expectation value) of O from

$$\langle O(T) \rangle = \frac{\sum_C O(C) e^{-E(C)/kT}}{\sum_C e^{-E(C)/kT}} = \sum_C O(C) \frac{e^{-E(C)/kT}}{Z} = \sum_C O(C) w(C) \tag{4.33}$$

C represents a configuration of the Ising variables and the summation is over the set of all possible configurations. In an application, O might represent the energy, the magnetization or the spin-spin correlation function. But this is possible only in very small systems, because in order to compute for each configuration the probability $w(C)$, we have to evaluate the partition function, which requires the computation of all the Boltzmann factors. For larger systems the best one can do is to average over a subset of the states. They are chosen randomly from some probability distribution ρ. In this way choosing M states the average becomes

$$\langle O \rangle = \frac{\sum_i^M O_{\mu_i} \rho_{\mu_i}^{-1} w_{\mu_i}}{\sum_j^M \rho_{\mu_j}^{-1} w_{\mu_j}} \tag{4.34}$$

As the number of states M increases it becomes more and more accurate in estimating the value of $\langle O \rangle$. The problem is to find a good probability distribution to be able to find a good representative subset. If one knows which states make the most important contribution to the observable, and if we can pick up only those states, we would get a much better estimate of O. This is the basic idea of Importance sampling.
We have better to assume the probability of each state as the probability of randomly pick up

exactly that state. So the the estimator fo $\langle O \rangle$ takes the form

$$O_M = \frac{1}{M} \sum_i^M O_{\mu_i} \tag{4.35}$$

This form is much useful especial if the system spends most of its time in few states, they will be picked up with an higher frequency. If all the measurements of O are statistically independent, the estimate of the statistical error σ_O associated with this average is

$$\sigma_O \approx \sqrt{\frac{1}{M-1}\left(\frac{1}{M}\sum_{i=1}^M O_{\mu_i}^2 - \left(\frac{1}{M}\sum_{i=1}^M O_{\mu_i}\right)^2\right)} \tag{4.36}$$

Now how exactly we pick our states so that one appears with the correct probability needs a deep analysis using Markov process.

Let us move now to a quantum system and so to quantum Monte Carlo method. What distinguishes a quantum Monte Carlo method from a classical one is the initial effort necessary to represent the quantum problem in a suitable form for simulation [70]. Almost always, the transformation replaces the quantum degrees of freedom by classical ones, and it is to these classical degrees of freedom that the Monte Carlo method is actually applied. In classical example, a configuration C or a state corresponds to one outcome $X = (Y_1, Y_2, \dots, Y_n, \dots)$ of the random variable in the sample space. In a quantum Monte Carlo simulation, we may have to distinguish between the state of the system and the Monte Carlo configuration. Moreover, to deal with the state in quantum mechanics, we need to project it on a basis set $|C\rangle$ of configuration states, then we can evaluate matrix elements $\langle C|H|C\rangle$ and wave functions $\psi(C) = \langle C|\psi\rangle$, which are classical objects. After that we can follow a Markov process as before.

4.3.1 Markov process

When one wants to perform a Monte Carlo simulation one needs a way to generate an appropriate random set of states according to the correct probability distribution. This can be done with a procedure called Markov processes [68].
It is a method useful to sample any density function regardless of analytic complexity in any number of dimensions, but it is only asymptotically accurate [65]. This sampling method has also the drawback of introducing strong correlation between consecutive random variables. This, if compared to independent samples, leads to an increase in variance for a certain number of steps.

When a systems is in equilibrium, it shows statistical properties independent on its kinetics. So we can leave the system free to evolve by stochastic transition, from one state μ to a new random one ν. The probability of evolving from state μ to state ν is called transition probability $P(\mu \to \nu)$. For a true Markov process all transition probabilities should be constant in time, and independent on any previous states

$$P(\mu \to \nu) = P(\nu|\mu) = P(\nu|\mu, \mu_{n-1}, \dots, \mu_1) \tag{4.37}$$

Moreover they should respect the constraint

$$\sum_\nu P(\mu \to \nu) = 1 \tag{4.38}$$

since the Markov process must generate some state every time. The state μ itself is included in the sum, representing no transition
The repetition of Markov process generate a Markov chain of states. That, if long enough, eventually produces a succession of states which follow the requested probability distribution ρ, whatever was the starting point. When this happens we will say that it comes to equilibrium, or that it is thermalized, since it is the analogue of a classical system that reaches the thermal equilibrium with its surrounding. In order to achieve this we place two further conditions on our Markov process: ergodicity and detailed balance.

Ergodicity

The condition of ergodicity is the requirement that, during our Markov process, any state of the system should be reachable from any other, if the procedure runs for long enough. This is necessary if we want to generate states according to one probability distribution. Every state appears with some non zero probability so it must be accessible from any initial state otherwise its probability will be zero. There must be always at least one path of non zero probability between any two states.

Detailed balance

Detailed balance ensures that when we reach equilibrium we have generated the exact distribution rather than any other. We can say that the system is in equilibrium when the rate at which it makes transitions into and out of any state μ are equal

$$\sum_{\nu} w_{\mu} P(\mu \to \nu) = \sum_{\nu} w_{\nu} P(\nu \to \mu) \tag{4.39}$$

This condition is called global balance. Since we know the sum is normalized, we can simplify the stationary condition as

$$w_{\mu} = \sum_{\nu} w_{\nu} P(\nu \to \mu) \tag{4.40}$$

When any set of transition probabilities satisfy this equation, the probability distribution w_{μ} is an equilibrium distribution. But this condition is not sufficient to obtain the correct limit after long time. Following [64] and [70], we can try to go deeper into this problem.
The transition probabilities $P(\mu \to \nu)$ can be thought as the elements of a matrix $\mathbf{P}$, called Markov matrix or the stochastic matrix for Markov process. Since we use $w_{\mu}(t)$ for probability that the system is in state μ at time t, after one step along our Markov chain (state ν at time $t+1$) the probability will be

$$w_{\nu}(t+1) = \sum_{\mu} P(\mu \to \nu) w_{\mu}(t) \tag{4.41}$$

or in matrix notation as

$$\mathbf{w}(t+1) = \mathbf{P} \cdot \mathbf{w}(t) \tag{4.42}$$

where $\mathbf{w}(t)$ is the vector of probabilities. The Markov process reaches equilibrium for $t \to \infty$ if the state satisfies

$$\mathbf{w}(\infty) = \mathbf{P} \cdot \mathbf{w}(\infty) \tag{4.43}$$

however it is possible to reach a condition of fictitious equilibrium, where the dynamics is confined in a limit cycle, a dynamic equilibrium. In this case the probability satisfies the condition

$$\mathbf{w}(\infty) = \mathbf{P}^{n} \cdot \mathbf{w}(\infty) \tag{4.44}$$

where n is the limit cycle's length. So the only condition of eq.(4.40) does not guarantee that the generated states will follow the predicted probability distribution, since eq.(4.40) could be fulfilled even when the process is confined in a limit cycle.
To overcome the problem one can apply an additional condition, the detailed balance

$$p_\mu P(\mu \to \nu) = p_\nu P(\nu \to \mu) \tag{4.45}$$

this is able to eliminate the limit cycles. It represents the equality of transition from one state to another and vice versa. On average the system should go from one state to another with the same probability as for the opposite transition. In a limit cycle this condition is violated, since some preferential direction is created. Once the limit cycles are removed the system will always tend to the correct probability distribution after a long time, as $t \to \infty$, $\mathbf{w}(t)$ will tend exponentially towards the eigenvector corresponding to the largest eigenvalue of $\mathbf{P}$.

Stochastic matrices

Let us justify now, on the basis of [70][75], the remarkable properties of a Markov chain and discuss the general conditions for their validity.
Our discussion is focused on the transition probability $\mathbf{P}$: if it is an ergodic stochastic matrix it allows a unique stationary distribution satisfying eq.(4.40). A stochastic matrix is a non-negative matrix where the sum of each column (or each row) is 1, $\sum_i P_{ij} = 1$.
A matrix is defined irreducible if there are no permutations of rows and columns that can lead the transformation into the form

$$\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \tag{4.46}$$

where A_1 and A_3 are square matrices of any possible order. If such a permutation exists, then the matrix is called reducible. Aperiodicity means that there are no permutations that transforms P_{ij} into the form

$$\begin{pmatrix} 0 & A_1 & 0 & \cdots & 0 \\ \vdots & 0 & A_2 & \ddots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & A_{n-1} \\ A_n & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.47}$$

An irreducible and aperiodic stochastic matrix is called an ergodic matrix. In general, an irreducible and aperiodic non negative matrix is called a primitive matrix. An ergodic matrix is a primitive stochastic matrix.
In general, a stochastic matrix is not symmetric, $P_{ij} \neq P_{ji}$. Consequently, it has unequal left (x^α) and right (y^α) eigenvectors that share the same eigenvalues

$$\sum_i x_i^\alpha P_{ij} = \lambda_\alpha x_j^\alpha, \ \sum_j P_{ij} y_j^\alpha = \lambda_\alpha y_i^\alpha \tag{4.48}$$

If the eigenvalues are distinct, $\lambda_\alpha \neq \lambda_\beta$, then the right and left eigenvectors, x^α and y^β, are linearly independent and satisfy $[x^\alpha]^T \cdot y^\beta = 0$. One can show that an irreducible stochastic matrix has a non-degenerate eigenvalue equal to 1 and the corresponding right eigenvector's components are all positive [75]. These results follow from the application of the Perron-Frobenius theorem [76] (Perron proved results for a positive matrix and then Frobenius extended them to non-negative matrices). We will call $\rho(A)$ spectral radius of a matrix A. It is the largest absolute value of any eigenvalue of A. If this matrix is real and non-negative, $\rho(A)$ is an eigenvalue

that corresponds to a non negative unique eigenvector (up to an overall scaling). If moreover this matrix is irreducible, then it has a non-degenerate eigenvalue equal to its spectral radius. Accordingly, this eigenvalue is real and positive.
If the matrix is stochastic, then we can show that the absolute value of the eigenvalue cannot exceed unity, and therefore at least one of the dominating eigenvalues is unity, and other dominating eigenvalues, if any, are complex numbers with absolute value 1.
Moreover, if we require the additional condition of aperiodicity, we can show that, if the matrix $\mathbf{P}$ is aperiodic as well as irreducible, there exists a finite number n of matrix multiplications so that $\mathbf{P}^n$ is a positive matrix; that is, $[P^n]_{ij} > 0$ for any pair of i and j. When a non-negative matrix satisfies this condition, it is called primitive, and the sampling is called ergodic [70]. An irreducible, stochastic matrix can have one ore more zero elements. These corresponds to forbidden transition for the process. But this matrix must have at least one non-zero off-diagonal element in each column. In general, the same is true for the product of a stochastic matrix with itself.

Meyer showed in [75] that, for a positive stochastic matrix, the eigenvector corresponding to eigenvalue equal to 1 is unique. So, it follows that, these properties also apply to the primitive stochastic matrix. Finally if $\mathbf{P}$ is chosen so that it is a primitive stochastic matrix and satisfy stationary condition (4.40) with a given target distribution p_μ, the right dominating eigenvector, corresponding to eigenvalue one, equals p_μ. It means that after sufficiently many iterations in the Markov-chain Monte Carlo simulation, the probability distribution converges to the target distribution.
Having more than one of this eigenvectors means that the system subdivide in mutually inaccessible regions, but if the ergodicity is satisfied then such subsets are forbidden and hence there in only one of this eigenvectors.

Now, let us take a vector ψ, it can be expressed as a linear combination of the above cited right eigenvectors y^α

$$\psi = \sum_\alpha a_\alpha y^\alpha \tag{4.49}$$

where we assume there is an initial overlap $a_1 = 1$ with the dominant eigenvector. We can order the eigenvalues as $1 > |\lambda_2| \geq |\lambda_3| \geq \ldots$ If now we multiply the matrix $\mathbf{P}$ k-times with this vector

$$P^k\psi = y^1 + \sum_{\alpha\geq 2} \lambda_\alpha^k a_\alpha y^\alpha \tag{4.50}$$

As k becomes large, the eigenvectors with non-dominant eigenvalues project out, leaving just the right eigenvector y^1. Setting $y = p > 0$ leads us to eq.(4.40). The vector of probabilities tends exponentially to the desired distribution, it is independent on the starting state p^0 provided that there is an initial overlap a_1, but we have to iterate for a sufficiently large k to reach equilibrium. This condition can be connected to any distribution p_μ, but we still have to select a set of transition probabilities which satisfy equation (4.45). Our constraint leave a lot of freedom over how to choose the transition probabilities.
When we write a code we have to respect the time interval needed for equilibration, only after a suitable long interval we can start to sample points for computing the expectation. However it could happen that high probability regions of phase space are connected by low probability paths. So even if the code is formally ergodic, it could behave as if it is not and spend much more time in some specific regions of space. In this case transition to different regions is a rare event, and the system can lock in a non representative phase. Such a situation requires special

care to ensure statistical independence among measurements.

4.3.2 Metropolis algorithm

We have discussed the Markov process' potential for achieving the equilibrium distribution, but it is not at all clear which Markov process is the most suitable for generating the required transition probabilities. While there are numerous alternative ways to transform an existing state into a new one, we might not find one that provides the precise set of transition probabilities. However we do not have to. Following Metropolis et al.[68]:
The purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules. Classical statistics is assumed ...
Even though the Metropolis algorithm was developed to address issues in classical equilibrium statistical mechanics, it quickly expanded to all stochastic process-based areas of study. It turns out that by adding an acceptance ratio, we may use whatever algorithm we choose to create the new states and still get the set of transition probabilities we want. The detailed balancing requirement is satisfied by the stochastic transition matrix that is defined. Let us analyse the sample from this matrix.
Since the condition of sum rule of probabilities leave some freedom in our process, one is free to modify the probability to go to different states $P(\mu \to \nu)$, and then one balances all using $P(\mu \to \mu)$ the probability to stay in same state. If we make an adjustment like this we have also to simultaneously making change in $P(\nu \to \mu)$, so their ratio and detailed balance are preserved. In practice we have enough freedom that the transition probabilities can take any set of values we like by changing the $P(\mu \to \nu)$. Let us split it in two

$$P(\mu \to \nu) = g(\mu \to \nu)A(\mu \to \nu) \tag{4.51}$$

The first quantity in eq.(4.51) is the selection probability, given an initial state it is the probability that the algorithm will generate a new target state ν. The matrix $\mathbf{g}$ must satisfy

$$g_{\mu\nu} \geq 0, \ \sum_{\mu} g_{\mu\nu} = 1 \tag{4.52}$$

and it must be a probability distribution that we are able to sample exactly.
The second term in eq.(4.51) is the acceptance ratio. If a test state is generated from an initial state, it should be accepted as the next state a fraction of the time equal to $A_{\mu\nu}$. When this happens, our system changes its state, otherwise it will be unchanged. The freedom in $\mathbf{A}$ reflects the freedom in $\mathbf{g}$, because the important part is the ratio

$$\frac{P(\mu \to \nu)}{P(\nu \to \mu)} = \frac{g(\mu \to \nu)A(\mu \to \nu)}{g(\nu \to \mu)A(\nu \to \mu)} \tag{4.53}$$

Also the sum rule is satisfied since the system must end up in some state after each step in Markov chain. So the algorithm generates a new state with some probability $\mathbf{g}$, that could be accepted or rejected with ratio $\mathbf{A}$. Low ratio means a waste of time, since the system will be rarely moved. If the ratio is too high the result is similar, the change in state is too small and the correlation is enormous. A chain of too similar of equal states is the principal source of statistical inefficiency since we are looking for independent measurements. Newman and Barkema [64] suggest to embody the dependence of $\mathbf{P}$ on characteristic of the states in the selection probability $\mathbf{g}$, in this way one can keep the acceptance ratio high as much as

possible. Newman and Barkema's ideal algorithm selects the new states with exactly the correct transition probability so they are always accepted.
The first justification of the algorithm was defined by Wood and Parker [77], in their publication the algorithm is connected to a Markov process. After them the justification got stronger with the work by Rosenbluth [78] which is discussed in appendix C.

4.3.3 Implementation of the algorithm

Since now the algorithm is solid and justified, Let us develop its implementation for a future application in a simulation. We can start supposing that the system is in a random state μ. We know that whatever the initial state is, it has no influence on our result once the system thermalizes. So let us assume that the system has now a probability distribution of states f.
To reach the equilibrium a transition to some sate ν is proposed using any distribution $\mathbf{g}$ from μ, then one has to compare $f(\nu)$ with $f(\mu)$ and, taking into account $\mathbf{g}$, the system can be moved to ν, accepting the transition, or the system can stay in μ, rejecting the transition. The acceptance occurs with probability $A(\mu \to \nu)$, which must be calculated so as to satisfy detailed balance

$$A(\mu \to \nu)g(\mu \to \nu)f(\mu) = A(\nu \to \mu)g(\nu \to \mu)f(\nu) \tag{4.54}$$

we expect that both $\mathbf{g}$ and f will play a significant role in determining $\mathbf{A}$.
On the basis of what we said in previous sections, if we take a set of n steps in a random walk, $\nu_1, \ldots, \nu_n$, where the random variable ν is a many-dimensional vector. Whichever is the pdf $\phi_i(\nu)$ associated to such a variable, we expect a convergence in distribution of the pdf

$$\lim_{n\to\infty} \phi_n(\nu) = f(\nu) \tag{4.55}$$

At each step in the random walk we assume the possibility of transition in both directions $\mu \leftrightarrow \nu$, and, following [65] we can define the quantity $q(\mu \to \nu)$ as

$$q(\mu \to \nu) = \frac{g(\nu \to \mu)f(\nu)}{g(\mu \to \nu)f(\mu)} \geq 0 \tag{4.56}$$

and then compute the probability of accepting a move as

$$A(\mu \to \nu) = \min(1,\ q(\mu \to \nu)) \tag{4.57}$$

this is the Metropolis-Hasting algorithm, a generalization by Hasting [79]. For a symmetrical $\mathbf{g}$, one reduces to the original Metropolis algorithm, where $\mathbf{A}$ depends just on ratio of probability of states. In case of classical system it is directly connected to the evolution towards minimal energy. However, in both cases, we do not need a normalized stationary distribution.
In concrete form, the algorithm starts with the random variable ν which assumes, after n steps of the random walk, the value ν_n. Then one has to consider a new step along the walk. one can find ν'_{n+1} as possible successive value, it is sampled from $g(\nu_n \to \nu'_{n+1})$, according to previous equation the probability $q(\nu_n \to \nu'_{n+1})$ of accepting ν'_{n+1} is computed.
Now we have two possibilities in eq.(4.57): if $q > 1$ then the transition is accepted and the system moves to new state ν_{n+1}, if $q < 1$, then $A(\nu_n \to \nu'_{n+1}) = q(\nu_n \to \nu'_{n+1})$, so there is no certainty in changing state. Here a new random variable ξ appears, it is exploited to define a random threshold to accept or reject the transition. One has that if $A(\nu_n \to \nu'_{n+1}) > \xi$, then $\nu_{n+1} = \nu'_{n+1}$; otherwise $\nu_{n+1} = \nu_n$, in this case the same state as before is taken into account for sampling and extracting the average. We expect that the walk tends to lock in a region of

real minimum, encountering high probabilities barrier, so very small transition probabilities **A**, that leads to large contribution in final estimate. However relative minima can affect this ideal behaviour, so we have to take care in choosing how two states are connected. If they are too far, just a small local minima is able to stop the walk. However, too close states can also make it difficult to cross barriers and cause extremely long simulations. Moreover, these two limiting cases result in extremely strong correlations.
When we complete the number of steps we consider as enough we can stop the walk. If we are interested in a quantity $O(\nu)$, that, from sampled points $\nu_1, \ldots, \nu_N$, assumed values $O_1, \ldots, O_N$, the estimate of observable's expectation is approximated as

$$\langle O \rangle = \int O(\nu) f(\nu) d\nu \approx \frac{1}{N} \sum_{i>j}^{N+j} O(\nu_i) \tag{4.58}$$

Where j is the last discarded element. If we were able to sample uncorrelated points, according to central limit theorem, we are allowed to assume a normal distribution of the observable and associate an uncertainty given by one standard deviation (we are satisfied with the 68% confidence interval)

$$\sigma_O \approx \sqrt{\frac{1}{N-1} \left(\frac{1}{N} \sum_{i>j}^{N+j} O_i^2 - \left(\frac{1}{N} \sum_{i>j}^{N+j} O_i \right)^2 \right)} \tag{4.59}$$

Because $f(\nu)$ is sampled only asymptotic the estimate of the integral is biased. An amount that can be made smaller by discarding more and more ν_i at the beginning of the walk and by extending the walk to larger N.

Let us now check that the algorithm leads to the correct asymptotic distribution. In order to do this, we can analyse which is the relationship between distributions of two consecutive steps. If we assume that after n steps ot the random walk the distribution is $\phi_n(\nu)$, after one step there are two contributions: The probability of entering the vicinity of ν from another point μ, given by $g(\mu \to \nu)\phi_n(\mu)d\mu$, and the probability to be in ν and remain there. Weighting the first based on the probability that the move will be accepted, it becomes $A(\mu \to \nu)g(\mu \to \nu)\phi_n(\mu)d\mu$, so for any initial point μ the probability to reach ν is given by

$$\int A(\mu \to \nu) g(\mu \to \nu) \phi_n(\mu) d\mu \tag{4.60}$$

the probability of non-accepted movement away from ν is

$$\int (1 - A(\nu \to \mu)) g(\nu \to \mu) d\mu \tag{4.61}$$

So the relation with the distribution at the next step is

$$\phi_{n+1}(\nu) = \int A(\mu \to \nu) g(\mu \to \nu) \phi_n(\mu) d\mu + \phi_n(\nu) \int (1 - A(\nu \to \mu)) g(\nu \to \mu) d\mu \tag{4.62}$$

We expect that, after many iteration of this relation, we reach an asymptotic distribution $f(\nu)$. According to Feller [80], if a random walk defines a system that is ergodic, so there is not any periodic repetition of position in the neighbourhood of any initial point, then an asymptotic pdf exist and it is unique if

$$\phi_n(\nu) = f(\nu) \to \phi_{n+1} = f(\nu) \tag{4.63}$$

that is if $f(\nu)$ is a stationary point of recursion.We assume the proof or ergodicity of our algorithm [77]. Let us set $\phi_n(\nu) = f(\nu)$ in eq.(4.62), then the relation for ϕ_{n+1} becomes

$$\phi_{n+1}(\nu) = \int A(\mu \to \nu) g(\mu \to \nu) f(\mu) d\mu + \int (1 - A(\nu \to \mu)) g(\nu \to \mu) f(\nu) d\mu \tag{4.64}$$

because of validity of detailed balance and normalization of $\mathbf{g}$, the relation simplifies to

$$\phi_{n+1}(\nu) = \int g(\nu \to \mu) f(\nu) d\mu = f(\nu) \tag{4.65}$$

and $f(\nu)$ is guaranteed to be the asymptotic distribution of the random walk.

Even if we can reach the correct distribution $f(v)$ only asymptotically, this algorithm is endowed with great utility. In fact it enables us to sample from very complicated many-dimensional probability distribution functions in a straightforward way.
Aside from this general considerations, we have to adapt every time the algorithm to our specific problem since always a part of the sampled points has to be discarded, and the number must be estimated. Therefore, the better it is the proximity of $\phi_1(\nu)$ and $f(\nu)$, the smaller will be the waste, since we start the process nearer the asymptote. So it is of vital importance to define a correct and powerful way to find how long we must wait before starting sampling, and also how long in simulation steps is the correlation generated by the algorithm. This will be done in following subsection.

4.4 Variational Monte Carlo

Let us start discussing the first of two quantum Monte Carlo methods that will be encountered in this thesis. It is based on Monte Carlo sampling and the variational principle from quantum mechanics, so first a brief introduction to latter concept will be presented.

4.4.1 Variational principle

According to the variational principle, when we cannot find an analytic solution to Schrödinger equation, we are able to estimate the energy of the ground state of a system. When one considers a trial function, different from the real ground state, depending on some variational parameters, the corresponding energy eigenvalue will be greater than the minimal energy of the system. If we consider the normalized ground state wavefunction $|\Phi_0\rangle$, we have

$$\langle \Phi_0 | \Phi_0 \rangle = 1 \tag{4.66}$$

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E_0 \tag{4.67}$$

where $\hat{H}$ is the system Hamiltonian and E_0 is the ground state energy. If now we take one approximation of $|\Phi_0\rangle$, that depends on some parameters $|\Phi\rangle_\alpha$, it leads to the variational energy

$$E_V = \min_\alpha E[\alpha] = \min_\alpha \frac{\langle \Phi | \hat{H} | \Phi \rangle_\alpha}{\langle \Phi | \Phi \rangle_\alpha} \geq E_0 \tag{4.68}$$

Despite being omitted, the proof is included in every textbook on quantum mechanics.
This theorem implies a strategy to find the ground state of the system. If we are able to write a wave function dependent on some adjustable parameters, which we suppose can parametrize an

interesting subset of the Hilbert space, these parameters can be varied until the minimum energy is obtained. In general the variational space can be highly dimensional to increase accuracy and therefore the minimization procedure could lead to some difficulties. In this thesis we had to optimize only two parameters, so a simple grid in parameter space was needed.

4.4.2 Monte Carlo application

We can start from eq.(4.68), Let us write it in a first quantization form as

$$E_V = \min_\alpha \frac{\int d\mathbf{R}\Phi^*_\alpha(\mathbf{R})\langle\mathbf{R}|\hat{H}|\Phi\rangle_\alpha}{\int d\mathbf{R}\Phi^*_\alpha(\mathbf{R})\Phi_\alpha(\mathbf{R})} \tag{4.69}$$

Because of what we said in previous sections, we are interested in rewriting this integral as the expectation value of some quantity, in order to apply our knowledge of Monte Carlo. We call this observable local energy

$$E^\alpha_L(\mathbf{R}) = \frac{1}{\Phi_\alpha(\mathbf{R})}\langle\mathbf{R}|\hat{H}|\Phi\rangle_\alpha \tag{4.70}$$

So the variational energy becomes

$$E_V = \min_\alpha \frac{\int d\mathbf{R}|\Phi_\alpha(\mathbf{R})|^2 E^\alpha_L(\mathbf{R})}{\int d\mathbf{R}|\Phi_\alpha(\mathbf{R})|^2} = \int d\mathbf{R} w_\alpha(\mathbf{R}) E^\alpha_L(\mathbf{R}) \tag{4.71}$$

Now, we can see that the trial wavefunction must not be normalized and positive everywhere, because when we compute the probability

$$w_\alpha(\mathbf{R}) = \frac{|\Phi_\alpha(\mathbf{R})|^2}{\int d\mathbf{R}'|\Phi_\alpha(\mathbf{R}')|^2} \tag{4.72}$$

any constant cancels out, and the square protects against negative values. Moreover we do not have to calculate w. In fact, if we substitute this result in eq.(4.57), where we are assuming a final state in position $\mathbf{R}_f$ and an initial one in $\mathbf{R}_i$, we obtain

$$A(\mathbf{R}_f|\mathbf{R}_i) = \min\left[1, \frac{g(\mathbf{R}_i|\mathbf{R}_f)w(\mathbf{R}_f)}{g(\mathbf{R}_f|\mathbf{R}_i)w(\mathbf{R}_i)}\right] = \min\left[1, \frac{g(\mathbf{R}_i|\mathbf{R}_f)|\Phi_\alpha(\mathbf{R}_f)|^2}{g(\mathbf{R}_f|\mathbf{R}_i)|\Phi_\alpha(\mathbf{R}_i)|^2}\right] \tag{4.73}$$

Then one can evaluate the transition from $\mathbf{R}_i$ to $\mathbf{R}_f$ and accept or reject this movement according to previous equation. The process is then continued iteratively.
We still have no prescriptions to define the selection probability. It can assume a wide variety of forms, depending on specific problem some of them can reduce the correlation of successive samples. First one can assume a symmetric g, thus simplifying further eq.(4.73). However, Hastings proved that using a nonsymmetric g, the algorithm's efficiency can be significantly enhanced [79]. If the transition probability is ergodic, as discussed before, we can reach the equilibrium distribution after a sufficiently large number of iterations.
The simplest algorithm one can create is the one in which every new step begins with the addition of some fixed quantity δ multiplied by a vector of random numbers with the same dimension of $\mathbf{R}$ and with each component $\xi_i \in [-1, 1]$

$$\mathbf{R}_f = \boldsymbol{\xi}\delta + \mathbf{R}_i \tag{4.74}$$

One better option is to replace the displacement with a Gaussian distributed one. Then probabilities and energies are computed. Typically one tries to set the quantity added in such a

way that in average the fraction of accepted moves is between 40 - 60%. After each step, but only when equilibrium is reached, E_L^α is collected and stored. At the end of the process the average of all sampled energies leads to the best estimate with an associated uncertainty, which is estimated with the variance of E_L^α.

When the process is concluded we have at our disposal a set of energies depending on some parameters. If one is able to identify the exact solution of the Schrödinger equation, as for example in case of harmonic oscillator, the energy is the exact ground state. Otherwise, when we know only some information about the quantum system, if we trust the ergodicity of our algorithm, we can extract the minimal energy from our data. We assume it is a good approximation of the ground state, but it may be greater of the real minimum.
However, we are able to obtain an estimate of the energy level at equilibrium. We are assuming that the state of the system does not change during the process, the evolution is performed by some fictitious components, collective variables representing a point in configuration space [81], the so called walkers. They move in a fictitious space according to transition probability, looking for positions of maximal wavefunction amplitude. At the end their distribution defines the average energy. This explains why ergodicity is so important to obtain a correct result, we need to be able to inspect all the space and avoid stopping in local minima.
Therefore, the number of walkers has some importance in simulation. However In Variational Monte Carlo each walker is independent, and increasing their number corresponds to enlarge the statistic of the simulation, so the same effect of a longer simulation time if the algorithm is ergodic.

4.4.3 Smart Monte Carlo

One possibility to improve the variational Monte Carlo method is the introduction of a drift velocity. In stead of the time independent Schrödinger equation one starts from the Fokker-Plank equation for diffusion characterized by a time-dependent density $f(\mathbf{R}, t)$. Following [82] [83] one can find the corresponding Langevin equation that leads to a modification of the transition matrix. In addition to the random part, there is also a time-dependent drift component in updating the state

$$\mathbf{R}_f = \mathbf{R}_i + D\mathbf{F}(\mathbf{R}_i)t + \boldsymbol{\chi} \tag{4.75}$$

where D is the diffusion constant, $\boldsymbol{\chi}$ is a gaussian random vector with zero mean and variance $2Dt$. This change increases the acceptance with respect to standard VMC, and the convergence is faster, but at the price of an higher computational cost. In this thesis we will maintain the Variational method as discussed at the beginning of the section.

4.5 Diffusion Monte Carlo

Let us move now to Diffusion Monte Carlo, where we consider the time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi(\mathbf{R}, t)}{\partial t} = \hat{H}\Psi(\mathbf{R}, t) = -D\nabla^2\Psi(\mathbf{R}, t) + \hat{V}\Psi(\mathbf{R}, t) \tag{4.76}$$

where the potential $\hat{V}$ can include both interaction and external potentials, and we introduced the notation $D = \hbar^2/2m$. In contrast to previous method, we are now trying to solve the problem exploiting the time evolution of the quantum state. Now, following [84], Let us introduce

an imaginary time $\tau = -it/\hbar$. Then, since a constant can be added to the potential without changing the form of the Ψ, we shift the Schrödinger equation by an energy E, whose meaning will become clear later:

$$-\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \hat{H}\Psi(\mathbf{R},\tau) = -D\nabla^2\Psi(\mathbf{R},t) + (V-E)\Psi(\mathbf{R},t) \tag{4.77}$$

We can expand a formal solution of eq.(4.77) in eigenstates of $\hat{H}$

$$\Psi(\mathbf{R},\tau) = e^{-\tau(\hat{H}-E)}\Psi(\mathbf{R},0) = e^{-\tau(\hat{H}-E)}\sum_i c_i\phi_i(\mathbf{R},\tau) = \sum_i c_i\phi_i(\mathbf{R},0)e^{-(E_i-E)\tau} \tag{4.78}$$

where $\mathbf{R}$ is a 3N-dimensional vector, with N the number of particle of the system, ϕ_n and E_n eigenfunctions and eigenenergies of the problem, the latter ordered as $E_0 < E_1 < \ldots$. The amplitudes of the components in eq.(4.78) change with time, either increasing or decreasing, depending on the sign of $(E_i - E)$. At large times the term that corresponds to the projection on the ground state dominates the sum. In other words all excited states decay exponentially fast and only contribution from ground state survives

$$\Psi(\mathbf{R},\tau) = c_0\phi_0(\mathbf{R},0)e^{-(E_0-E)\tau} \text{ for } \tau\to\infty \tag{4.79}$$

The solution then transforms into a sum of exponentially decaying terms after being moved to imaginary time. Thus instead of a time-dependent superposition of states, imaginary time acts as a projector that, at large times, projects onto the lowest energy state.

The formal solution of the Schrödinger equation written in coordinate space is given by

$$\langle\mathbf{R}|\Psi(\tau)\rangle = \sum_{\mathbf{R}'}\langle\mathbf{R}|e^{-\tau(\hat{H}-E)}|\mathbf{R}'\rangle\langle\mathbf{R}'|\Psi(0)\rangle \tag{4.80}$$

or in terms of Green's function

$$\Psi(\mathbf{R},\tau) = \int G(\mathbf{R},\mathbf{R}',\tau)\Psi(\mathbf{R}',0)d\mathbf{R}' \tag{4.81}$$

that one can interpret as a probability of moving the set of walkers from $\mathbf{R}$ to $\mathbf{R}'$ in time τ. In other words, the differential Schrödinger equation (4.77) corresponds to the integral equation (4.81), which can be integrated with help of Monte Carlo methods.
The problem is that one in general does not know the exact G, but the use of Monte Carlo is still possible by means of short-time approximation, solving eq.(4.81) step by step

$$\Psi(\mathbf{R},\tau+\Delta\tau) = \int G(\mathbf{R},\mathbf{R}',\Delta\tau)\Psi(\mathbf{R}',\tau)d\mathbf{R}' \tag{4.82}$$

In this approximation we can write an explicit analytical form for G. Then we have to introduce a source of error neglecting the non-commutativity of $\hat{T}$ and $\hat{V}$. If completely removed the error order is $O(\tau^2)$, and the resulting short-time approximation equation is

$$G(\mathbf{R},\mathbf{R}',\delta\tau) = \langle\mathbf{R}'|e^{-\delta\tau V}e^{-\delta\tau T}|\mathbf{R}\rangle e^{+\delta\tau E} + O(\delta\tau^2) \tag{4.83}$$

otherwise one could do a little better and reach order $O(\delta\tau^3)$, in this case we have

$$\begin{aligned} G(\mathbf{R},\mathbf{R}',\delta\tau) &= \langle\mathbf{R}'|e^{-\delta\tau V/2}e^{-\delta\tau T}e^{-\delta\tau V/2}|\mathbf{R}\rangle e^{+\delta\tau E} + O(\delta\tau^3) \\ &= \exp\left[-\delta\tau\frac{V(\mathbf{R}')+V(\mathbf{R})}{2} + \delta\tau E\right]\langle\mathbf{R}'|e^{-\delta\tau T}|\mathbf{R}\rangle + O(\delta\tau^3) \end{aligned} \tag{4.84}$$

Equations (4.83) and (4.84) can be rewritten in this approximation as

$$G(\mathbf{R}, \mathbf{R}', \delta\tau) = G_D(\mathbf{R}, \mathbf{R}', \delta\tau) w(\mathbf{R}, \mathbf{R}'; \delta\tau) \tag{4.85}$$

where $w(\mathbf{R}, \mathbf{R}'; \delta\tau)$ is called the branching, it contains the V dependence. In the first case it is given by

$$w(\mathbf{R}, \mathbf{R}'; \delta\tau) = \exp\left[\delta\tau(E - V)\right] \tag{4.86}$$

while in eq.(4.84) it is

$$w(\mathbf{R}, \mathbf{R}'; \delta\tau) = \exp\left[-\delta\tau \frac{V(\mathbf{R}') + V(\mathbf{R})}{2} + \delta\tau E\right] \tag{4.87}$$

The propagator $G_D = \exp[-\delta\tau T]$ is simply the solution for ordinary diffusion:

$$G_D(\mathbf{R}, \mathbf{R}', \delta\tau) = (4\pi D\delta\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R})^2}{4D\delta\tau}\right] \tag{4.88}$$

In both cases $\delta\tau$ must be small for the validity of the approximation. When we iterate the code, one step will be equivalent to the integral

$$\Psi(\mathbf{R}, \tau) = \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}', \delta\tau) \Psi(\mathbf{R}', \tau - \delta\tau) \tag{4.89}$$

Unfortunately, the above procedure can be very inefficient, due in part to $w(\mathbf{R}, \mathbf{R}'; \delta\tau)$, that defines how walkers are multiplied or cancelled. When the potential becomes large and negative, the number of copies of a random walker near this singularity would become exceedingly large, leading to the collapse of the simulation. In general, efficiency can be improved when known information is incorporated into a simulation.
Kalos and coworkers [85] found an effective importance sampling scheme. They introduced information about the system through a trial wave function $\Psi_T(\mathbf{R})$. In this method the imaginary time Schrödinger equation is expressed in terms of the product $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau)$ and is transformed as

$$-\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = -D\nabla^2 f(\mathbf{R}, \tau) + D\nabla \cdot [f(\mathbf{R}, \tau) F(\mathbf{R})] + [E_L(\mathbf{R}) - E] f(\mathbf{R}, \tau) = \tilde{H} f \tag{4.90}$$

where

$$\begin{aligned} E_L &= \frac{H\Psi_T}{\Psi_T} \\ F &= \frac{2}{\Psi_T(\mathbf{R})} \nabla \Psi_T(\mathbf{R}) \end{aligned} \tag{4.91}$$

are the local energy and the drift quantum force. We can recognize the analogy with classical system, where $p(\mathbf{R}) \propto \exp(-U(\mathbf{R}))$, if we approximate a quantum probability distribution by the square of a trial wave function, since the force is an antigradient of the potential energy $F = -\nabla U(\mathbf{R}) = \nabla \ln p(\mathbf{R})$.
In particular eq.(4.90) resembles a diffusion plus a branching process, but now there is a drift velocity imposed on the diffusion. The singularities in the branching factor are now almost gone replacing V with E_L. In fact, E_L approaches a constant as $\Psi_T \to \phi_0$. Furthermore, the drift term now guides the walk preferentially into regions of space where the wave function is large. Following [84][86] we can write (4.90) for the Green's function as

$$-\frac{d\tilde{G}}{d\tau} = \tilde{H}\tilde{G} = (\tilde{T} + \tilde{V})\tilde{G} \tag{4.92}$$

with

$$\begin{aligned}\tilde{T} &= -D\nabla^2 + D\nabla \cdot F + DF \cdot \nabla \\ \tilde{V} &= E_L - E\end{aligned} \tag{4.93}$$

and rewrite the Green's function as

$$\tilde{G}(\mathbf{R}, \mathbf{R}', \delta\tau) = \tilde{w}(\mathbf{R}, \mathbf{R}'; \delta\tau)\tilde{G}_D(r \to \mathbf{R}',\ \delta\tau) \tag{4.94}$$

where $\tilde{w}$ is the same as before but replacing V with E_L, while $\tilde{G}_D$ is the propagator for diffusion with a drift and is solution to (4.90) without the branching term

$$\tilde{G}_D(\mathbf{R}, \mathbf{R}', \delta\tau) = \langle \mathbf{R}'|e^{-\delta\tau\tilde{T}}|\mathbf{R}\rangle = (4\pi D\delta\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R}' - \mathbf{R} - D\delta\tau F(\mathbf{R}))^2}{4D\delta\tau}\right] \tag{4.95}$$

We see that the importance sampling procedure is essentially the same as the original method, but now after each step we perform two more calculations E_L and F, they are needed to evaluate $\tilde{w}$ and $\tilde{G}_D$, respectively. The modified procedure now updates the product function f

$$f(\mathbf{R}, \tau) = \int d\mathbf{R}'\tilde{G}(\mathbf{R}, \mathbf{R}', \delta\tau)f(\mathbf{R}', \tau - \delta\tau) \tag{4.96}$$

It could be convenient to split the right hand of eq.(4.90) in three terms, as $\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_3$

$$\begin{aligned}\hat{H}_1 &= -D\nabla^2 \\ \hat{H}_2 &= D\left((\nabla \cdot F) + F \cdot \nabla\right) \\ \hat{H}_3 &= E_L(\mathbf{R}) - E\end{aligned} \tag{4.97}$$

This, rewritten in the coordinate representation, gives the expression for the Green's function

$$\tilde{G}(\mathbf{R}, \mathbf{R}', \tau) = \iint G_1(\mathbf{R}, \mathbf{R}_1, \tau)G_2(\mathbf{R}_1, \mathbf{R}_2, \tau)G_3(\mathbf{R}_2, \mathbf{R}', \tau)d\mathbf{R}_1 d\mathbf{R}_2 \tag{4.98}$$

The Green's function should satisfy Bloch differential equation:

$$\begin{cases} -\frac{\partial}{\partial\tau}G_i(\mathbf{R}, \mathbf{R}', \tau) = \hat{H}_i G_i(\mathbf{R}, \mathbf{R}', \tau), \ \text{ for } i = 1, 2, 3 \\ G_i(\mathbf{R}, \mathbf{R}', 0) = \delta(\mathbf{R} - \mathbf{R}') \end{cases} \tag{4.99}$$

The equation for the kinetic term has the form of a diffusion equation, which solution is a Gaussian

$$G_1(\mathbf{R}, \mathbf{R}', \tau) = (4\pi D\tau)^{-3N/2} \exp\left[\frac{(\mathbf{R}' - \mathbf{R})^2}{4D\tau}\right] \tag{4.100}$$

The equation for the drift force term has solution

$$G_2(\mathbf{R}, \mathbf{R}', \tau) = \delta(\mathbf{R}' - \mathbf{R}(\tau)) \tag{4.101}$$

here $\mathbf{R}(\tau)$ is the solution of the classical equation of motion

$$\begin{cases} \frac{d\mathbf{R}(\tau)}{d\tau} = DF(\mathbf{R}(\tau)) \\ \mathbf{R}(0) = \mathbf{R} \end{cases} \tag{4.102}$$

The last equation describes the branching term and has solution

$$G_3(\mathbf{R}, \mathbf{R}', \tau) = \exp\left[(E - E_L(\mathbf{R}))\tau\right] \delta(\mathbf{R} - \mathbf{R}') \tag{4.103}$$

In this case of importance sampling, the Green's function contains a non Hermitian operator $\tilde{T}$, so we have to impose detailed balance to guarantee equilibrium, accepting the move of walker from x to y with Metropolis probability

$$A(y, x; \delta\tau) = \min(1, q(y, x; \delta\tau)) \tag{4.104}$$

where

$$q(y, x; \delta\tau) = \frac{|\Psi(y)|^2}{|\Psi(x)|^2} \frac{\tilde{G}(x, y; \delta\tau)}{\tilde{G}(y, x; \delta\tau)} \tag{4.105}$$

However we can exploit the $\delta t = 0$ extrapolation, that leads to the ground state eigenfunction and so we can extract a correct result.

4.5.1 Implementation of the method

In previous section the basic theoretical aspects of the method have been introduced, now they will be reformulated in a more suitable form for coding.
If the function $f(\mathbf{R}, \tau)$ is real and positive, as it happens in case of ground state of a bosonic system, it can be treated as population density distribution and $f(\mathbf{R}, \tau)d\mathbf{R}$ gives the probability to find a walker at time τ in the vicinity $d\mathbf{R}$ of a point $\mathbf{R}$. In terms of Markov chains the Green's function $G(\mathbf{R}, \mathbf{R}', \tau)$ is the transition matrix which determines the evolution of the distribution. Let us now discuss the tasks that must be implemented in a code following [84]. First of all we have to create an initial random configuration, representing the function $f(\mathbf{R}, t = 0)$. We know that the final result is independent on that choice, however it is convenient to start from the final configuration of a Variation Monte Carlo simulation when it is possible. This initial state will be diffused over a large number of small time intervals, every time we sample the observables. We know also that there is a certain amount of iteration necessary to arrive at convergence, during this process we must skip the sampling, then the distribution of walkers will be near the optimal one. When we perform one step $\tau \to \tau + \Delta\tau$, we have to take into account the contribution of each walker, and every time evaluate each term in (4.95). The first means the diffusion of each of the walkers in the configuration space

$$\mathbf{R}^{(1)}(\tau + \Delta\tau) = \mathbf{R}(\tau) + \boldsymbol{\xi} \tag{4.106}$$

where ξ is a normally distributed random vector with standard deviation $\sigma = \sqrt{2D\Delta\tau}$, here we need the Box-Muller method for sampling. The second term describes the action of the drift force, which guides the walkers to places in the configuration space, where the trial wave function is maximal. This is the way how importance sampling acts in the algorithm

$$\mathbf{R}^{(2)}(\tau + \Delta\tau) = \mathbf{R}(\tau) + DF(\mathbf{R})\Delta\tau \tag{4.107}$$

Since the Green's functions of steps (4.100),(4.101) are normalized to unity, the normalization of wave function f is conserved and then the number of walkers remains constant. So we can express both steps as

$$\mathbf{R}'(\tau + \Delta\tau) = \mathbf{R}(\tau) + DF(\mathbf{R})\Delta\tau + \boldsymbol{\xi} \tag{4.108}$$

The third term is different, since the corresponding Green's function is no longer normalized, it changes the population of walkers. We call it the branching term, and it acts as

$$f^{(3)}(\tau + \Delta\tau) = \exp\left[(E - E_L(\mathbf{R}))\Delta\tau\right] f(\mathbf{R}, \tau) \tag{4.109}$$

The weight of a walker $\mathbf{R}$ changes in relation to its local energy, and it is expressed as a number of possible replicas for each walker. This is computed based on its weight as

$$n = [b + \eta] = [\exp\{\Delta\tau \left(E_L(\mathbf{R})/2 + E_L(\mathbf{R}')/2 - E\right)\} + \eta] \tag{4.110}$$

it is the integer part of the exponential b plus a random number $\eta \sim U(0, 1)$. Then the walker is duplicated $[b + \eta]$ times, with a fixed maximum of copies in order to avoid divergence in population. If $n = 0$ we delete the walker. Now it is clear that by adjusting the value of E as

$$E = E + \alpha \ln\left(\frac{\text{ideal number of walkers}}{\text{actual number of walkers}}\right) \tag{4.111}$$

where α is a small number, one can control the size of the population and keep it within the desired range. The branching is an essential part of the Diffusion Monte Carlo algorithm, unlike the other two terms it corrects the trial wave function, and in case of an exact eigenfunction of the Hamiltonian its action is simply the identity. In fact in this case the local energy is independent on $\mathbf{R}$, thus the branching becomes irrelevant, since it will be just a constant.
In conclusion, as in Variational Monte Carlo, we extract our best estimate of an observable as average from sampled quantities with associated relative uncertainty. Instead, since we have introduced the short-time approximation, we have a bias in each result that depends on the time interval. Therefore, in order to extract the non-approximated result valid only for $\Delta\tau = 0$, the same simulation will be repeated for different time steps. After verifying that we are in a linear regime for the time-step bias and we can project to the correct energy by means of a linear regression.

4.5.2 Fixed-node diffusion Monte Carlo method

The Fixed-node Diffusion Monte Carlo (FN-DMC) method modifies our previous concepts, which are correct for bosons, in order to allow an approximate treatment of fermionic systems, where we have to take into account the antisymmetry of the wavefunction. As consequence the wavefunction in some regions in configuration space is negative. Neglecting this effect will lead us to simulate a bosonic system. We will call the regions where the wavefunction vanishes the nodal surface. It is a hyper-surfaces that divides the configuration space into connected hyper-volumes called nodal pockets.
In [70] they define this as the sign problem:
The exponential increase of the Monte Carlo errors with increasing system size or decreasing temperature that often accompanies a Markov chain simulation whose limiting distribution is not everywhere positive.
Many approaches to this problem have been proposed (some different strategies are presented in [83]), even if a general solution is not possible, as demonstrated in [87]. When we know the exact location of the nodal surface, the region in configuration space that delineates the change in sign, we can use sign-weighted averages [70]. However in general we do not have this knowledge, so we have to prevent the error scaling. To do this we will follow [88] using the importance sampling method, in particular we will approximate the nodal surface of the true wavefunction with that of a trial function. In this way it will be possible to consider $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau)$ always positive and therefore recover a density as in the bosonic case. Moreover we add an infinite repulsive potential that forbids walkers' movements across the nodal surface defining impenetrable barriers. Every time a walker attempts to cross one of these boundaries, it is killed, and its replication number is set to zero.
Therefore in the FN-DMC method we solve the many body Schrödinger equation in a series

of non-intersecting domains subjected to Dirichlet boundary conditions. In a simulation the function $f(\mathbf{R}, \tau)$ is evolved in imaginary time according to eq.(4.90) with addition of a potential. Thus, the FN-DMC energy depends on a good parametrization of the many-body nodal surface. Hammond et al. [83] showed by means of perturbation theory argument, that the first non-zero energy correction is second order, so the dependence is not crucial. Moreover, Reynolds et al. [88], shown that, due to the nodal constraint, the fixed-node energy is a variational upper bound to the exact eigenenergy for a given symmetry. One possible shortcoming of the method is the additional time bias, which raises from the possibility of a double cross of the node if the time step is sufficiently large.

4.6 Finding equilibrium and correlation

Let us now discuss how we can find a meaningful estimate for the length of simulation before two points can be approximated as independent. Since we are going to generate highly correlated points, neglecting of this fact will always cause an underestimated statistical error because the central limit theorem for the variance requires that the sampled points are independent. We will adopt the blocking technique as in [81] [70] [89], but also many others. The method consists in separating the sampled points in a number N_b of blocks, each with $M = N/N_b$ points. The block average is

$$\bar{x}_b = \frac{1}{M} \sum_{i}^{M} x_i \tag{4.112}$$

This new variable has the same expectation value of the original x_i random variable. Then the global average, that is our best estimate, is

$$\bar{x} = \frac{1}{N_b} \sum_{b}^{N_b} \bar{x}_b \tag{4.113}$$

It is a random variable with the same expectation, but now, if the length of the blocks is sufficiently large so that their averages $\bar{x}_b$ are independent, we expect it can follow the central limit theorem, so finally compute the variance as in eq.(4.36) (where one has to replace M with N_b).
We need a way to define which is the best block dimension and their best number. In [89] they use the autocorrelation time

$$T_c = 1 + \frac{2}{\mathrm{VAR}[x] M} \sum_{k<l} \mathrm{Cov}[x_k, x_l] = M \frac{\mathrm{VAR}[\bar{x}_b]}{\mathrm{VAR}[x]} \tag{4.114}$$

that is equal to 1 in absence of correlation. They impose as condition for a correct estimation of statistical uncertainty $M > 100 T_c$. Since this process requires the computation of a correlation between successive steps, so a large CPU demand and therefore a long iteration time, we prefer to estimate this decorrelation phase following [81], where it is analysed the trend of the estimated error with respect to block size to find a convergence in number of points per block. One has to run the code for a large number of iterations. Then this set of acquired data must be averaged changing every time the block dimension. As one can see in figure 4.1 for the blue triangles, the error trend has an initial increase from very small values, when blocks of too few points are implemented. The error continued to raise until at a certain block size, if the number of sampled points is sufficiently large, it stabilizes and converges to some value. When this happens we can be pretty sure that the system has reached an equilibrium distribution,

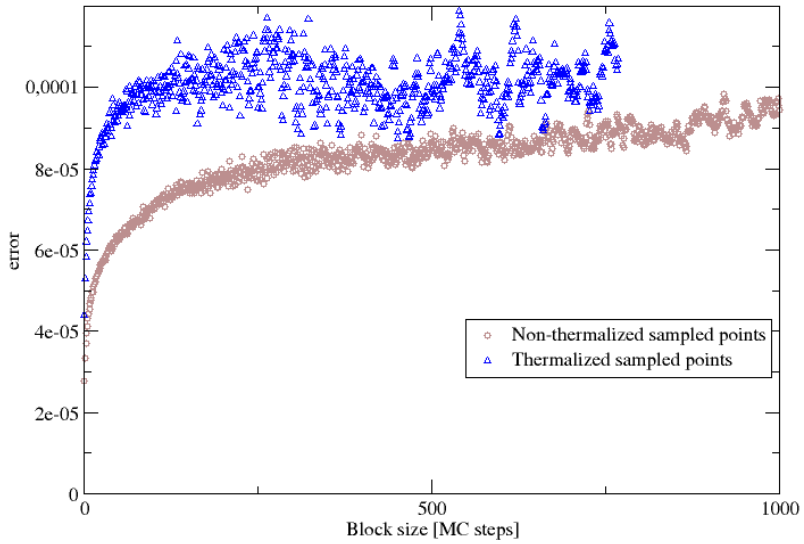


Figure 4.1: Plot showing the various errors' behaviours (for the average of the sampled points) depending on the convergence of the distribution. When the system reach equilibrium (thermalization), the sampled points fluctuate normally around their average, and the associated error assumes a well-defined behaviour as explained in the text. One can observe an increasing trend in the case of zero discarded points.

so we can extract a minimum number of points per block. We will consider this number as the size to have independent variables, and for the total number of points we multiply our result by $N_b \approx 50$, the number of blocks for a reasonable final average. Subsequently, the number of iterations can be increased for a lower statistical error. As discussed before, there is a transient phase between the beginning of the simulation and the equilibrium, where points are not sampled. If one forgets this phase the result is the set of brown circles, in this case it is required a much larger block size. In order to fulfil this condition, at least the first block will be discarded every time.

4.7 Wavefunction

Until now we have underlined many times the importance of the trial wavefunction for a correctly approximated result, but we have said nothing about how to construct this function. There are many possibilities in literature, one can see [83], and we will follow [81]. Depending on the work dedicated to crafting the wavefunction, one can go more or less closer to the real ground state. Knowing the symmetry of the problem could mean an increase of the parametrization of the trial state and a reduction of variational energy.
Since we are interested in dilute systems, if the system was composed of bosons the Jastrow wavefunction [63][90] would be sufficient to capture few-body correlations. In our case we deal, however, with fermions, so we need to anti-symmetrize the wavefunction and create a nodal surface. We will consider the most simple solution, which consists in a Slater determinant, which proved to be very accurate [91]. Another possibility is the BCS determinant [92][93], which we avoid since we are interested in the normal phase of a Fermi gas.

In constructing our trial function we must try to use all information which is available. Moreover we have to set appropriate boundary conditions to take into account the effect of the short-range two-body interactions when two fermions are close to each other. We call these asymptotic relations cusp conditions, which also leads to an enforcement of the sampling accuracy [81]. They are expressed in three dimension as:

$$\lim_{r_i \to r_j} \Psi(r_1, ..., r_i, ..., r_j, ..., r_N) = \left(\frac{1}{|r_i - r_j|} - \frac{1}{a} \right) \phi \left[\frac{r_i + r_j}{2}, r_k, k \neq i, j \right] \tag{4.115}$$

valid for $|r_i - r_j|$ greater than the potential range R, and a is the scattering length. In our analysis of fermionic gas with two spin populations, we will use the product of a Jastrow function and two Slater determinants

$$\Psi^{JS} = \Phi_J D_\uparrow D_\downarrow \tag{4.116}$$

4.7.1 Jastrow function

The first term in eq.(4.116) is the symmetrized product of a few-body function, its simplest form contains only two-body terms

$$\Phi_J(R) = \prod_{i<j} f(r_{ij}) \tag{4.117}$$

where the product includes all couple of interacting particles (i, j), with the notation $r_{ij} = |\vec{r}_i - \vec{r}_j|$. If short range approximation of the interaction is valid, $f(r_{ij})$ can be the solution of two-body problem. Then it automatically satisfies the cusp conditions. In case of two spin types $(\uparrow, \downarrow)$:

$$\Phi_J(R) = J_{\uparrow\uparrow}(R) J_{\uparrow\downarrow}(R) J_{\downarrow\downarrow}(R) = \prod_{i<j} f_{\uparrow\uparrow}(r_{ij}) \prod_{i,a} f_{\uparrow\downarrow}(r_{ia}) \prod_{a<b} f_{\downarrow\downarrow}(r_{ab}) \tag{4.118}$$

4.7.2 Slater determinant

The second and third components of our wavefunction eq.(4.116) solve the problem of antisymmetry and related nodal surface. The simplest possibility is a couple of Slater determinants constructed adiabatically from non-interacting wavefunctions $\phi_{k_p}(r_i)$, since we are assuming the validity of Fermi Liquid theory that will be discussed in the next chapter. Since we are in dilute conditions, we will use orbitals of single non-interacting particles corresponding to eigenvectors of momentum in a box with periodic boundary conditions. In [93][61] they are proved to be enough to have a comparison with experiments. So we use the determinant D

$$D_\sigma(R) = \det S^\sigma_{pi} = \det \left[\phi_{k^\sigma_p}(r_i) \right] \tag{4.119}$$

where single orbitals are expressed as exponentials, that can be combined in sine and cosine functions, a more convenient form in performing numerical calculation because they allow for using a purely real wavefunction.

4.7.3 Energy calculation

After we presented the wavefunction, we are now interested in computing the the kinetic energy term K of the local energy

$$E_L(\mathbf{R}) = \frac{\langle \mathbf{R}|\hat{H}|\Psi^{JS}\rangle}{\Psi^{JS}(\mathbf{R})} \tag{4.120}$$

where $\mathbf{R}$ is a point in configuration space. The goal is to compute explicitly the quantum force components that appears in K, that in representation coordinate is the sum of all second derivatives for all particles for both spin species:

$$K = -\frac{D}{\Psi^{JS}}\left(\sum_i \sum_\alpha \partial^2_{\alpha_i}\Psi^{JS} + \sum_a \sum_\alpha \partial^2_{\alpha_a}\Psi^{JS}\right) \tag{4.121}$$

where the constant $D = \hbar^2/2m$, the masses of all particles are assumed equal, and alpha runs on the three coordinates $\alpha = x, y, z$. following [81], after some calculations reported in appendix D, we obtain

$$\begin{aligned} K^{JS} &= K^J + K^S_\uparrow + K^S_\downarrow - \frac{D}{2}\left[\sum_{i\,\alpha} F_i^{J\,\alpha}\cdot F_i^{S\,\alpha} + \sum_{a\,\alpha} F_a^{J\,\alpha}\cdot F_a^{S\,\alpha}\right] \\ &= K^J + K^S_\uparrow + K^S_\downarrow - \frac{D}{2}\mathbf{F}^J\cdot\mathbf{F}^S \end{aligned} \tag{4.122}$$

where

$$\begin{aligned} K^J &= -\frac{D}{\Phi_J}\left(\sum_i\sum_\alpha \partial^2_{\alpha_i}\Phi_J + \sum_a\sum_\alpha \partial^2_{\alpha_a}\Phi_J\right) = \\ &-\frac{D}{4}\mathbf{F}_{\uparrow\uparrow}\cdot\mathbf{F}_{\uparrow\uparrow} + D\sum_{i,\,j\neq i}\left[e^L_{\uparrow\uparrow}(r_{ij}) + \left(\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\right)^2\right] - \frac{D}{4}\mathbf{F}_{\downarrow\downarrow}\cdot\mathbf{F}_{\downarrow\downarrow} + D\sum_{a,\,b\neq a}\left[e^L_{\downarrow\downarrow}(r_{ab}) + \left(\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\right)^2\right] \\ &-\frac{D}{4}\mathbf{F}_A\cdot\mathbf{F}_A + 2D\sum_{i,a}\left[e^L_{\uparrow\downarrow}(r_{ia}) + \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2\right] - \frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A = K_{\uparrow\uparrow} + K_{\uparrow\downarrow} + K_{\downarrow\downarrow} - \frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A \end{aligned} \tag{4.123}$$

$$\mathbf{F}_S = \left\{F^\alpha_{i\uparrow\uparrow},\, F^\alpha_{a\downarrow\downarrow}\right\},\ \mathbf{F}_A = \left\{F^\alpha_{i\uparrow\downarrow},\, F^\alpha_{a\uparrow\downarrow}\right\}$$

$$\begin{aligned} K^S_\uparrow &= -\frac{D}{D_\uparrow}\sum_i\sum_\alpha^d \partial^2_{\alpha_i} D_\uparrow \\ &= -D\sum_i\sum_\alpha^d\left[\sum_{pq}\bar{S}_{qi}\partial_{\alpha_i}S_{iq}\bar{S}_{pi}\partial_{\alpha_i}S_{ip} - \sum_{pq}\bar{S}_{qi}\bar{S}_{pi}\partial_{\alpha_i}S_{iq}\partial_{\alpha_i}S_{ip} + \sum_p \bar{S}_{pi}\partial^2_{\alpha_i}S_{ip}\right] \\ &= -D\sum_i\sum_\alpha^d\sum_p \bar{S}_{pi}\partial^2_{\alpha_i}S_{ip} \end{aligned} \tag{4.124}$$

$$
\begin{aligned}
K_{\downarrow}^{S} &= -\frac{D}{D_{\downarrow}} \sum_{a} \sum_{\alpha}^{d} \partial_{\alpha_a}^2 D_{\downarrow} \\
&= -D \sum_{a} \sum_{\alpha}^{d} \left[\sum_{pq} \bar{S}_{qa} \partial_{\alpha_a} S_{aq} \bar{S}_{pa} \partial_{\alpha_a} S_{ap} - \sum_{pq} \bar{S}_{qa} \bar{S}_{pa} \partial_{\alpha_a} S_{aq} \partial_{\alpha_a} S_{ap} + \sum_{p} \bar{S}_{pa} \partial_{\alpha_a}^2 S_{ap} \right] \\
&= -D \sum_{a} \sum_{\alpha}^{d} \sum_{p} \bar{S}_{pa} \partial_{\alpha_a}^2 S_{ap}
\end{aligned}
\tag{4.125}
$$

CHAPTER 5

SIMULATION SETUP

In this chapter we want to present the details of how the physical system can be set in a computational feasible manner. So how the general notions on Monte Carlo method can be applied to solve our physical problem. In particular we want to describe the trial wave function and also explain how we computed the average energy and its units. Moreover, we will discuss also how one can choose important parameters like the number of walkers and the simulation length. The program for running a Monte Carlo simulation of this system was written by Dr. Gianluca Bertaina [1], and the work for this thesis did not include modifications to his code.

5.1 General aspects and units of the code

In all the simulations that will be carried out, we will consider a fermionic gas with two spin components, with respective number of atoms $N_\uparrow$ and $N_\downarrow$, both with the same mass m, in a three-dimensional (3D) box of size L. Because of low energy and high diluteness conditions, fermions with parallel spins will interact as free-particles, while atoms with antiparallel spins will feel a (3D) spherical well interaction potential with a range R, which will be much smaller than the size of the box. We will set periodic boundary conditions at the edge of the box, so we have to smoothly connect the scattering wavefunction in L. We will hold in every simulations the spin-$\uparrow$ particles as the majority component, and we will use their quantities as a reference. In particular we will hold fixed their density $\mathrm{n}_\uparrow$, then the total density $\mathrm{n}_{\uparrow\downarrow}$ will be rescaled with addition of minority particles as

$$\mathrm{n}_{\uparrow\downarrow} = \frac{N_\uparrow + N_\downarrow}{N_\uparrow} \times \mathrm{n}_\uparrow \tag{5.1}$$

In this way the box size will be always the same moving from case $N_\downarrow = 0$ to case $N_\downarrow > 0$. Instead a bigger box will be considered when increasing $N_\uparrow$.

We will limit the maximum number of successive rejected proposal moves by walkers to ten in the variational case and to six for diffusion Monte Carlo. This prevents one walker from being caught in some configuration; after the limiting rejection threshold is reached, the proposed move is forcefully accepted (thus exceptionally violating detailed balance law). In diffusion Monte Carlo, in addition, we will also set a maximum number of possible duplications (number of sons) equal to five. We expect that this limitation avoids high production of new identical

[1] Istituto Nazionale di Ricerca Metrologica, Strada delle Cacce 91, I-10135 Torino, Italy

walkers, which could slow down the simulation and introduce bias into the results, or in some cases lead to the breakdown of the simulation. We of course monitor the occurrence of these exceptions in the Monte Carlo algorithms, and check that their rate is smaller than 10^{-3} (typically less than 10^{-4})

In order to describe the units of the simulation results, let us first introduce the fundamental quantities of the code. They are:

- The reference length: R_{ref}.
- The reference mass: m_{ref}.
- The reference energy: E_{ref}.

Then the time interval dt will be in units of $[\hbar/E_{\text{ref}}]$. In the VMC simulations its value will be chosen to have an acceptance ratio around 50%, while for the DMC simulations we will start from half of the value used in VMC simulations and then we will decrease it further to extrapolate the limit $dt \to 0$.
The energy resulting from one simulation will be in units of reference energy E_{ref}, and rescaled by the total number of atoms and the Fermi energy (in units of E_{ref}) of the majority species

$$E_{QMC} = \frac{E\left[E_{\text{ref}}\right]}{N E_F^{\uparrow}[E_{\text{ref}}]} \tag{5.2}$$

It is convenient to use the Fermi energy of the majority species in the absence of interaction as the reference energy $E_{\text{ref}} = E_F^{\uparrow}$, and the fermion mass m as the reference mass. We want to keep fixed the quantities

$$\begin{aligned} E_{\text{ref}} &= \frac{\hbar^2}{2m_{\text{ref}}R_{\text{ref}}^2} = E_F^{\uparrow} \\ n_{\uparrow}R^3 &= 10^{-6} \end{aligned} \tag{5.3}$$

The last equation in (5.3) ensures that the physical range of the model potential has a negligible role, while the first equation yields

$$\frac{\hbar^2}{2m_{\text{ref}}R_{\text{ref}}^2} = \frac{\hbar^2}{2m_{\text{ref}}}(6\pi^2 n_{\uparrow})^{2/3} \to (6\pi^2 n_{\uparrow}R_{\text{ref}}^3)^{2/3} = 1 \to n_{\uparrow}R_{\text{ref}}^3 = \frac{1}{6\pi^2} \tag{5.4}$$

Let us substitute here $n_{\uparrow} = (k_F^{\uparrow})^3/(6\pi^2)$, then we obtain $R_{\text{ref}} = 1/k_F^{\uparrow}$. In conclusion our reference quantities are:

- The reference length: $R_{\text{ref}} = 1/k_F^{\uparrow}$
- The reference mass: $m_{\text{ref}} = m$
- The reference energy: $E_{\text{ref}} = E_F^{\uparrow}$

5.2 Trial wavefunction

Let us now explain how the wavefunction described in section 4.7 will be treated by the code. As a numerical expedient, the code works mainly with the logarithm of the functions, in this

way it is possible to reduce to comparable orders of magnitude the products of functions which could present large differences in magnitude.
In order to study the atomic system using a QMC simulation, we must first find a trial wave function that could be capable of including the properties of the exact but unknown wavefunction as best as possible. As we said before, we will use the product of a Jastrow function and two Slater Determinants

$$\Psi^{JS} = \Phi_J D_\uparrow D_\downarrow \tag{5.5}$$

as done in [61] [93] [51] [94], which proved to be an effective combination for investigation of the ultracold atomic gases in the normal phase.

5.2.1 Jastrow function

Let us separate the Jastrow function based on the different combinations of interacting spin pairs

$$\Phi_J(\mathbf{R}) = J_{\uparrow\uparrow}(\mathbf{R}) J_{\uparrow\downarrow}(\mathbf{R}) J_{\downarrow\downarrow}(\mathbf{R}) = \prod_{i<j} f_{\uparrow\uparrow}(r_{ij}) \prod_{i,a} f_{\uparrow\downarrow}(r_{ia}) \prod_{a<b} f_{\downarrow\downarrow}(r_{ab}) \tag{5.6}$$

Since atoms within the same spin population are not interacting, we can set the functions $f_{\uparrow\uparrow}$ and $f_{\downarrow\downarrow}$ as constants. Instead, the function $f_{\uparrow\downarrow}$ strongly depends on the relative position of the two atoms.
We have four important length scales: the s-wave scattering length, in our regime of interest it is divergent; the potential range R, that must be small since we are in condition of short-range potential; the average inter-particle distance $l = n^{-1/3}$ (in this case the highest density of the two species), which must be much bigger than R to motivate the use of the two-body function; the size of the simulation box L, which we assumed much bigger than l, to reduce finite-size effects. Moreover, there is also a variational parameter $\bar{R}$, which defines the relative distances after which we can match the two-body function $f_{\uparrow\downarrow}$ to an arbitrary smooth function. Obviously it must be greater than R.
When two atoms have a relative distance shorter than the potential range $r_{ia} < R$, we can use the solution of the radial Schrödinger equation just outside the well in the limit of vanishing energy for the two-body scattering problem

$$f_{\uparrow\downarrow}(r_{ia}) = A \frac{\sin(Z_0 r_{ia})}{r_{ia}} \tag{5.7}$$

where the parameter Z_0 depends on the potential depth V_0 as

$$Z_0^2 = \left(\frac{V_0 m}{\hbar^2} \right) , \tag{5.8}$$

and is set so as to correspond to a divergent scattering length via the equation (3.52) as explained in section 3.2.2.
Instead, when the atoms increase their relative distance, going over the potential range $R < r_{ia}$, since we are interested in conditions of resonance, we can simplify our function as

$$f_{\uparrow\downarrow}(r_{ia}) = \frac{1}{r_{ia}} \tag{5.9}$$

Equation (5.7) and (5.9) must be smoothly connected in $r_{ia} = R$, then, by enforcing the continuity of the functions and also of their derivative as in eq.(3.30), one is able to extract the

coefficient A, which is

$$A = \frac{1}{\sin(Z_0 R)} \ . \tag{5.10}$$

When the two atoms separate further we lose interest in an accurate description of their scattering wavefunction. So, within the interval of relative distances $\bar{R} < r_{ia} < L/2$, one can write

$$f_{\uparrow\downarrow}(r_{ia}) = c_1 + c_2 \left(e^{-\alpha r_{ia}} + e^{-\alpha(L-r_{ia})}\right) \ , \tag{5.11}$$

this function will tend to some constant at a distance equal to half the size of the box. Again the two regimes must be smoothly connected, then enforcing the same conditions of continuity for functions eq.(5.9)(5.11), one can determine the coefficients c_1 and c_2. They are

$$\begin{aligned} c_2 =& -\frac{1}{R^2}\frac{1}{-\alpha e^{-\alpha R} + \alpha e^{-\alpha(L-R)}} \\ c_1 =& \frac{1}{R} - c_2 \left(e^{-\alpha R} + e^{-\alpha(L-R)}\right) \end{aligned} \tag{5.12}$$

Because of the structure of $f_{\uparrow\downarrow}$ in eq.(5.11), its derivative is automatically vanishing for two atoms at the edge of the box $r_{ia} = L/2$. Moreover, it is useful that the limiting value of the wavefunction in eq.(5.11) is different from zero, so that there is a finite probability to have at least one fermion elsewhere inside the box. This allows for more freedom in the motion of the fermions during the simulations. Then, one can see we are left with a Jastrow function that has two variational parameters $\{\bar{R}, \alpha\}$ that we have to minimize with some preliminary variational Monte Carlo simulations.

Minimization of variational parameters

It is important to note that the first step for all subsequent simulations is to optimise the variational parameters. The best solution would be a parameter search algorithm that minimizes the energy within a given convergence range. However, due to the nature of the function, which has only two parameters, and for the level of precision required on the Monte Carlo average energies, it will be sufficient to proceed manually by extracting the parameters from the QMC results as the minima of a fitted second degree polynomial. After a few iterations, the parameters remain within their error range, and further iterations do not produce greater benefits on the average energies, within the required precision.
Furthermore, we noticed that these variational parameters have a weak dependence on the polarization of the gas (they are quite constant in particular when the number of atoms is around two hundred or larger). So when we inspect a wide range of polarization, we perform the minimization procedure only for some configurations and then we compute the other parameters by using linear interpolation.
Even though not perfectly accurate parameters do not rule out the possibility that the simulation converges to the correct value in the DMC method, they will increase the fluctuations around this result, forcing one to increase the estimated error for given simulation time. This fact can be demonstrated using the 1D harmonic oscillator model, for which we know the exact solution. In this case the Hamiltonian (in dimensionless units) is

$$H\psi(x) = \left[-\frac{1}{2}\frac{d}{dx} + \frac{1}{2}x^2\right]\psi(x) \tag{5.13}$$

with ground-state energy equal 1, and one can use the general trial function

$$\psi_T(x) = \exp(-\alpha x^2) \tag{5.14}$$

where α is a parameter fixed at the beginning of the simulation. As one can see in Tab.(5.1), even if $\alpha \neq 1/2$ one has the possibility to reach the correct energy for long enough simulations. However, the farther we are from the exact 1/2 coefficient, the longer it will take for the

Gaussian Coefficient α	Energy	Error
0.3	1.00000	0.00016
0.35	0.99993	0.00015
0.4	1.00010	0.00010
0.5	1	0
0.6	1.00010	0.00010
0.65	1.0000	0.0002

Table 5.1: Results for the ground state energy of a one dimensional harmonic oscillator obtained with the DMC method using different Gaussian trial wavefunctions with fixed coefficient α. For each case, the energy is obtained with a linear extrapolation to $dt \to 0$ of data obtained for six different time steps.

simulation to converge. Then, this phase of parameter optimization is important to reduce the variance of the results.

5.2.2 Slater determinant

We can start by considering a Slater determinant composed by orbitals of single non-interacting particles $\phi_{\mathbf{k}_p}(\mathbf{r}_i)$, where the columns are defined by the i-th particle while the momentum $\mathbf{k}_p$ defines the row of the determinant. The momenta chosen are those compatible with the simulation box, resulting from wavelengths λ such that L/λ is integer, where L is the length of the sides of the cubic simulation box. The smallest $\mathbf{k}_p$ will correspond to one oscillation $\lambda_{\max} = L$, while the others will be a multiple of this one.
These wavevectors can be expressed in terms of the vector $\mathbf{n} = (n_x, n_y, n_z)$ as $\mathbf{k} = \mathbf{n}2\pi/L$. The vector $\mathbf{n}$ leads to the quantum number $n^2_{\text{shell}} = n_x^2 + n_y^2 + n_z^2$, and we will say that the shell is closed if the fermions occupy all the possible states corresponding to wavevectors with the same number n^2_{shell} up to the number $n^2_{\text{closed shell}}$ included.
For a number N_σ of fermions of a given species, we fill progressively all available momentum states of increasing energy. Whenever possible, we insure that states with opposite momentum are either simultaneously occupied or left empty. In this way the total momentum of the system is zero. This can be achieved only when N_σ is odd (considering that the first fermion is placed in the zero momentum state). If the number of fermions is even then the momentum of the last fermion cannot be compensated and the system assumes a finite momentum.
The Slater Determinant will be treated in such a way as to transform the exponentials in sine and cosine functions. When N_σ is odd, our determinant becomes

$$D(\mathbf{x}_1, \ldots, \mathbf{x}_N) \propto \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \cos(\mathbf{k}_1 \cdot \mathbf{x}_1) & \cos(\mathbf{k}_1 \cdot \mathbf{x}_2) & \cdots & \cos(\mathbf{k}_1 \cdot \mathbf{x}_N) \\ \sin(\mathbf{k}_1 \cdot \mathbf{x}_1) & \sin(\mathbf{k}_1 \cdot \mathbf{x}_2) & \cdots & \sin(\mathbf{k}_1 \cdot \mathbf{x}_N) \\ \vdots & \vdots & \vdots & \vdots \\ \sin(\mathbf{k}_{n-1} \cdot \mathbf{x}_1) & \sin(\mathbf{k}_{n-1} \cdot \mathbf{x}_2) & \cdots & \sin(\mathbf{k}_{n-1} \cdot \mathbf{x}_N) \\ \cos(\mathbf{k}_n \cdot \mathbf{x}_1) & \cos(\mathbf{k}_n \cdot \mathbf{x}_2) & \cdots & \cos(\mathbf{k}_n \cdot \mathbf{x}_N) \\ \sin(\mathbf{k}_n \cdot \mathbf{x}_1) & \sin(\mathbf{k}_n \cdot \mathbf{x}_2) & \cdots & \sin(\mathbf{k}_n \cdot \mathbf{x}_N) \end{vmatrix} \tag{5.15}$$

otherwise when N_σ is even the last line will be given by cosine functions. In both cases, all the elements of the first line of the determinant are equal to 1 since $k_0 = 0$.
When one is computing the dispersion relation, and N_σ is even, additional complications arise from the combination of pre-existing system momentum and the new momentum caused by adding or eliminating one fermion. Then the total momentum is their vectorial sum, and the wavefunction is cumbersome to treat numerically. As a result of all of these complications, we opted to avoid an initial configuration with even N_σ when calculating dispersion relations.

Fixed Nodes for Diffusion Monte Carlo

In the case of the DMC, it was explained in section 4.5.2 that a nodal surface must be imposed to prevent negative values of the wave function. This is not necessary, instead, in VMC, since it uses the modulus squared of the wave function to proceed in the simulation.
In our case, the Slater determinant allows us to construct the nodal surface, where the wave function vanishes, as well as providing the wave function with the appropriate antisymmetry. The resulting nodal surface separates regions of space where the function does not change sign, then the algorithm behaves in these pockets as if the system was bosonic. The wavefunction will be forced to avoid sign changes by the code; this introduces a bias in the result that can be eliminated for time step approaching zero. Theoretically, this restriction would not be necessary for a very brief time step because, on the surface, the repulsive forces become infinite and would be sufficient to push away walkers approaching the crossing. The probability of crossing the surface increases as the time step is increased. Then in order to relax this approximation, and the one necessary to solve the diffusion equation, one has to perform many simulations for different dt, then the best estimate of DMC result will be the extrapolation to $dt = 0$. In our DMC simulations we will always perform at least six simulations for different values of dt.
Because of the impossibility to reach sufficiently small time steps, especially when increasing the number of atoms, we were forced to implement a cut-off for energies and forces [95]. This threshold depends on dt, it prevents the simulation from taking into account energies and forces that are too large in modulus, with the restriction to act only for values significantly larger than the typical energy variance, and at a low rate (we kept its order smaller than 10^{-4}). So the accepted energies E are those in the range

$$E_{\text{ref}} - \frac{E_{\text{cutoff}}}{\sqrt{dt}} < E < E_{\text{ref}} + \frac{E_{\text{cutoff}}}{\sqrt{dt}} \tag{5.16}$$

where one can set as reference the variational result.

5.3 Optimal system parameters

Let us discuss now how we can set the parameters of the simulation of the code. In particular we are interested in how to estimate the correct number of Monte Carlo steps, and the target number of walkers to reduce finite population errors.

5.3.1 Number of iterations

In order to estimate the length of the simulation we will rely mainly on a code that performs the block analysis of the variance as described in section 4.6. This program looks for convergence in the error's trend with respect to block size. It carries out a linear fit in the regime of block lengths where the variance reaches a plateau and verifies that the increase in the fitted variance

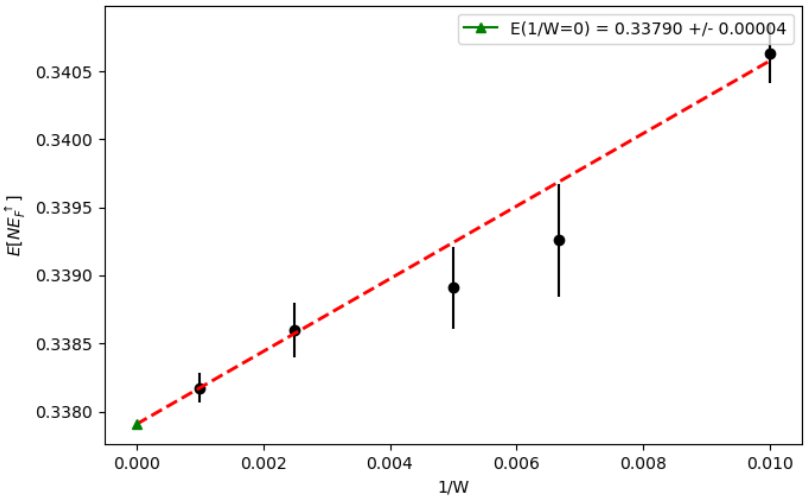


Figure 5.1: Extrapolation of the energy as a function of the inverse number of walkers W.

from the beginnig to the end of the plateau is smaller than 5%. Then, in order to prevent correlations between blocks, the code will set the minimum number of steps for convergence in each block. However, we will try to have always at least fifty blocks, a minimum number to obtain a meaningful average. Furthermore, the number of iterations will be further increased until the desired statistical error threshold is reached.

5.3.2 Number of walkers

The number of walkers plays a different role depending on the type of simulation. The variational simulations are unaffected by the number of walkers: in principle one can use just one walker, and the number of walkers has only the property of enlarging the sampled set of points. The situation is different for diffusion processes, where this number becomes quite important. In fact it affects the effects of branching and how well the reference energy is computed. If this number is too small the reference energy can be affected by fluctuations and imprint a drift towards non-minima regions. Following an analysis of the system's energy dependence on the number of walkers, as shown in figure 5.1, we made the choice to set this value to 1000. This will result in an additional systematic error of the order of 2.5×10^{-4}. This error must be summed in quadrature to the random error associated with the mean.

5.4 Finite size corrections

Since we are interested in analysing a very large system to approach the realization of the thermodynamic limit, the best we can do with our finite resources is to settle for a small system and then try to correct the finite size approximation. For two non-interacting systems of N_F and N'_F fermions, with the same density n, their energy can be connected as

$$\frac{E_{N'_F}}{N'_F} = \frac{E_{N_F}}{N_F} + \frac{T_{N'_F}}{N'_F} - \frac{T_{N_F}}{N_F} \tag{5.17}$$

where

$$\frac{T_{N_F}}{N_F} = \frac{1}{N_F} \sum_{i}^{\text{occ. orb.}} \frac{\hbar^2 k_i^2}{2m} \tag{5.18}$$

Without interactions eq.(5.17 equation is a tautology because the energy is exactly the kinetic energy. Let us assume the thermodynamic limit for the second system $N'_F \to \infty$, and, expressing the energy per particle as $e = E/N$ and $t = T/N$, the previous relation eq.(5.17) becomes

$$\begin{aligned} e_\infty &= e_{N_F} + t_\infty - t_{N_F} \\ t_\infty &= \frac{3}{5} E_F = \frac{3}{5} \frac{\hbar^2}{2m} \left(6\pi^2 n\right)^{2/3} \\ \frac{e_\infty}{E_F} &= \frac{e_{N_F}}{E_F} + \left(\frac{t_\infty}{E_F} - \frac{t_{N_F}}{E_F} \right) = \frac{e_{N_F}}{E_F} + \left(\frac{3}{5} - \frac{t_{N_F}}{E_F} \right) \end{aligned} \tag{5.19}$$

For interacting systems, if they are sufficiently large, by exploiting Fermi liquid theory and assuming that the effective mass is $m^* \sim m$, we can use the same finite size correction. In principle there is also a N_F-dependent correction, which we neglect for sufficiently large N_F, so we rewrite the last of eqs.(5.19) as

$$\frac{e_\infty}{E_F} = \frac{e_{N_F}}{E_F} + \frac{m}{m^*} \left(\frac{3}{5} - \frac{t_{N_F}}{E_F} \right) + c N_F^{-\nu} \sim \frac{e_{N_F}}{E_F} + \left(\frac{3}{5} - \frac{t_{N_F}}{E_F} \right) \tag{5.20}$$

However, considering a two component system, and in this case, following the same considerations as before, one can write eq.(5.20) into a spin dependent similar form

$$\begin{aligned} \frac{e^\uparrow_\infty}{E^\uparrow_F} &= \frac{e_{N\uparrow}}{E^\uparrow_F} + \frac{m}{m^*_\uparrow} \left(\frac{3}{5} - \frac{t_{N\uparrow}}{E^\uparrow_F} \right) \\ \frac{e^\downarrow_\infty}{E^\downarrow_F} &= \frac{e_{N\downarrow}}{E^\downarrow_F} + \frac{m}{m^*_\downarrow} \left(\frac{3}{5} - \frac{t_{N\downarrow}}{E^\downarrow_F} \right) \end{aligned} \tag{5.21}$$

The total energy of the mixture is $E_\infty = N^\uparrow e^\uparrow_\infty + N^\downarrow e^\downarrow_\infty$, then

$$\begin{aligned} \frac{E_\infty}{N E^\uparrow_F} &= \frac{1}{N^\uparrow (1+x) E^\uparrow_F} \left(\frac{N^\uparrow E^\uparrow_F e^\uparrow_\infty}{E^\uparrow_F} + \frac{N^\downarrow E^\downarrow_F e^\downarrow_\infty}{E^\downarrow_F} \right) \\ &\simeq \frac{1}{N^\uparrow (1+x) E^\uparrow_F} \left[N^\uparrow E^\uparrow_F \left(\frac{e_{N\uparrow}}{E_F} + \frac{m}{m^*_\uparrow} \left(\frac{3}{5} - \frac{t_{N\uparrow}}{E^\uparrow_F} \right) \right) + N^\downarrow E^\downarrow_F \left(\frac{e_{N\downarrow}}{E^\downarrow_F} + \frac{m}{m^*_\downarrow} \left(\frac{3}{5} - \frac{t_{N\downarrow}}{E^\downarrow_F} \right) \right) \right] \\ &= E_{QMC} + \frac{1}{1+x} \frac{m}{m^*_\uparrow} \left(\frac{3}{5} - \frac{t_{N\uparrow}}{E^\uparrow_F} \right) + \frac{x}{1+x} \left(\frac{k^\downarrow_F}{k^\uparrow_F} \right)^2 \frac{m}{m^*_\downarrow} \left(\frac{3}{5} - \frac{t_{N\downarrow}}{E^\downarrow_F} \right) \\ &= E_{QMC} + \frac{1}{1+x} \frac{m}{m^*_\uparrow} \left(\frac{3}{5} - \frac{t_{N\uparrow}}{E^\uparrow_F} \right) + \frac{x^{5/3}}{1+x} \frac{m}{m^*_\downarrow} \left(\frac{3}{5} - \frac{t_{N\downarrow}}{E^\downarrow_F} \right) \end{aligned} \tag{5.22}$$

where E_{QMC} as defined by eq.(5.2) is equal to $\left(N^\uparrow e^\uparrow_N + N^\downarrow e^\downarrow_N \right) / N E^\uparrow_F$ and $x = \left(k^\downarrow_F / k^\uparrow_F \right)^3$. However, since in the above correction we again assume $m/m^*_\uparrow = m/m^*_\downarrow = 1$, our finite size correction includes only the effects related to kinetic energy, so possible effects due to strong interaction are not accurately included and we take into account this lack of knowledge by increasing the energy uncertainty as explained in the following subsection.

5.4.1 Propagation of m^* uncertainty

The approximation in considering $m^* \sim m$ has an uncertainty that must be taken into account when one associates an error to the final average energy. Then we have to propagate its effects through all calculations and, in the end, sum it in quadrature to the QMC estimate of the error. From eq.(5.22) one has

$$\frac{\delta E_\infty}{N E_F^\uparrow} = \sqrt{(\delta E_{QMC})^2 + \left(\left|\frac{1}{1+x}\left(\frac{3}{5} - \frac{t_{N\uparrow}}{E_F^\uparrow}\right)\right| \delta_\uparrow\right)^2 + \left(\left|\frac{x^{5/3}}{1+x}\left(\frac{3}{5} - \frac{t_{N\downarrow}}{E_F^\downarrow}\right)\right| \delta_\downarrow\right)^2} \tag{5.23}$$

$$\delta_\sigma = \frac{m}{m_\sigma^*(x)} \frac{\delta m_\sigma^*(x)}{m_\sigma^*(x)} \tag{5.24}$$

where, inside the square root, the first term is computed by the code while the second two terms will be set to some constant as a first approximation. In particular, we have chosen to set $\delta m^*/m^* = 0.3$.
As we will show in Chap 6, where we evaluate the effective masses with VMC or DMC, we never find $m^* \gg 1.3m$, so the above finite size correction and uncertainty estimate are consistent with those findings. Two remarks are however in order: we use the finite-size correction only when interested in the absolute values of energy or chemical potential (see, e.g. Fig 6.2), while we do not apply them when evaluating energy dispersions, where we are only interested in energy differences. As second remark, if a FFLO transition indeed occurs, then effective masses are expected to diverge, so that the above finite-size correction with $m^* \sim m$ is not correct because it should tend to zero.

CHAPTER 6

RESULTS AND DISCUSSION

In this chapter we will present and discuss our results. As first thing it will be presented a validation of our code by comparing the resulting energies with literature. Then it will be presented our strategy to extract the effective mass from the quantum Monte Carlo outcomes. We can validate this method by comparing our work with theory, for what we obtained regarding the ideal Fermi gas, and with literature results (experiments and simulations), for the Fermi polaron. After that, we show our results for the effective masses of the quasi-particles, for both spin components, in a polarized ultracold Fermi gas at unitarity.

6.1 Validation of the code

Firstly we want to make sure that the simulations are efficient, so we can compare their outcomes to the known results at unitarity for a single spin-$\downarrow$ impurity in a spin-$\uparrow$ bath, and also in a range of polarization.

6.1.1 Energy of the Fermi polaron

As a first check of the code we can extract the energy of the polaron, adding one single impurity in a fully polarized system of fermions in the same volume. The energy of the polaron is given by

$$\Delta E = (N_\uparrow + 1)E_{\mathrm{QMC}} - (N_\uparrow)E_0 \tag{6.1}$$

where E_0 is the energy of the fully polarized system, and in case of DMC, E_{QMC} stands for its value extrapolated to $dt = 0$. This has been repeated for different numbers of majority atoms, in order to have the possibility to linearly extrapolate our Monte Carlo results with the inverse of $N_\uparrow = 19$, 27, 33, 57, both for Variational and Diffusion methods. The first method leads to $E_p/E_F^\uparrow = -0.482 \pm 0.002$, a result which is still far from the best estimates of the polaron energy available in the literature. The DMC method leads instead to a thermodynamic limit energy $E_p/E_F^\uparrow = -0.591 \pm 0.006$ that is fully compatible with what we expected for an attractive polaron as explained in section 3.4.1. This result is reported in figure 6.1.

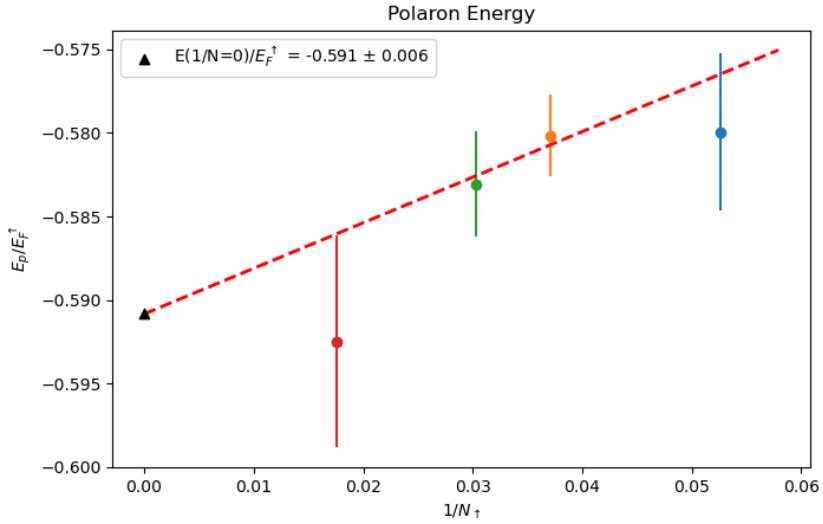


Figure 6.1: Thermodynamic limit extrapolation of the energy of the Fermi polaron by exploiting the Diffusion Monte Carlo method for $N_\uparrow = 19, 27, 33, 57$.

6.1.2 Energy of the polarized Fermi gas

In this case we will use the results by Nascimbène et al. [51] as our reference, and, in order to make this comparison, we have to change our units, moving from units of $[E_F^\uparrow N]$ to units of $[3E_F N/5]$. In a system composed by two spin species, with total number of atoms $N = N_\uparrow + N_\downarrow = N_\uparrow(1+x)$, these two units are connected by a multiplicative factor. So let us transform our energies reported in table 6.1 with the following equations

$$E_F^\uparrow = \frac{\hbar^2 k_F^{\uparrow 2}}{2m} = \frac{\hbar^2}{2m}\left(6\pi^2 n_\uparrow\right)^{2/3} = \frac{\hbar^2}{2m} 2^{2/3}\left(3\pi^2 n\right)^{2/3}(1+x)^{-2/3} \tag{6.2}$$

$$\frac{E}{NE_F^\uparrow} = \frac{E}{3NE_F/5}\frac{3}{5}\frac{E_F}{E_F^\uparrow} = E'\frac{3}{5}\left(\frac{k_F}{k_F^\uparrow}\right)^2 = E'\frac{3}{5}\frac{\left(3\pi^2 n\right)^{2/3}}{2^{2/3}\left(3\pi^2 n\right)^{2/3}}(1+x)^{2/3} = E'\frac{3}{5}\left(\frac{1+x}{2}\right)^{2/3} \tag{6.3}$$

Then one can see in figure 6.2 our results for both types of simulations (VMC and DMC) plotted together with the data extracted from [51], and one can conclude that we are obtaining reliable results. Then we are ready to start the computation of effective masses, so let us first introduce the method that will be exploited.

6.2 Dispersion relation

Let us now examine the relationship between the energies computed by QMC simulations and the dispersion relation.

6.2.1 Connection to simulation results

When one is interested in computing the dispersion relation, one needs three types of simulations (at fixed V), choosing N_σ as a closed shell while changing $N_{\bar{\sigma}}$ for considering different polarizations. The first one is for $E(N_\sigma, |\mathbf{k}| = k_F^{N_\sigma})$: this is a standard simulation that will act as reference for ground state energy. The second one is for $E(N_\sigma+1, \mathbf{k})$: here one adds a fermion

$N_\uparrow$	$N_\downarrow$	p	$E_{\rm VMC}[E_F^\uparrow N]$	$E_{\rm DMC}[E_F^\uparrow N]$
57	15	0.583333333	0.43803 ± 0.00012	0.41347 ± 0.00005
57	16	0.561643836	0.43078 ± 0.00018	0.40485 ± 0.00006
57	17	0.540540541	0.42346 ± 0.00015	0.39653 ± 0.00005
57	18	0.52	0.41661 ± 0.00018	0.38842 ± 0.00007
57	19	0.5	0.40981 ± 0.00020	0.38060 ± 0.00009
57	20	0.480519481	0.40550 ± 0.00017	0.37518 ± 0.00013
57	21	0.461538462	0.40177 ± 0.00016	0.36992 ± 0.00009
57	22	0.443037975	0.39764 ± 0.00016	0.36496 ± 0.00009
57	23	0.425	0.39393 ± 0.00023	0.36008 ± 0.00006
57	24	0.407407407	0.39013 ± 0.00020	0.35530 ± 0.00012
57	25	0.390243902	0.38625 ± 0.00021	0.35096 ± 0.00003
57	26	0.373493976	0.38328 ± 0.00016	0.34634 ± 0.00014
57	27	0.357142857	0.37977 ± 0.00019	0.34208 ± 0.00010
57	28	0.341176471	0.37871 ± 0.00016	0.33995 ± 0.00018
57	29	0.325581395	0.37742 ± 0.00018	0.33782 ± 0.00011

Table 6.1: Energies of a polarized Fermi gas as the polarization varies, at fixed volume. The polarization is selected on the basis of the number of atoms in the two populations, in this case the number of majority atoms is kept fixed at $N_\uparrow = 57$, while the number of minority atoms varies as shown in the second row of the table.

with momentum $|\mathbf{k}| > k_F^{N_\sigma}$ in a shell with $n^2_{\rm shell} > n^2_{\text{closed shell}}$. The last one is for $E(N_\sigma - 1, \mathbf{k})$: here one removes a fermion with momentum $|\mathbf{k}| \leq k_F^{N_\sigma}$ in a shell with $n^2_{\rm shell} \leq n^2_{\text{closed shell}}$.
Since the momentum $\mathbf{k}$ could be greater or smaller than $k_F^{N_\sigma}$, one has to calculate the dispersion relation in a different way for these two cases.
If one is interested only in the derivatives of $\varepsilon(\mathbf{k})$, then one can shift the dispersion by the chemical potential. This is valid for both spin components ($\sigma = \uparrow, \downarrow$) and leads to

$$\varepsilon_{N_\sigma}(\mathbf{k}) + \mu^{N_\sigma} = \begin{cases} E(N_\sigma + 1, \mathbf{k}) - E(N_\sigma, |\mathbf{k}| = k_F^{N_\sigma}) & \text{for} \quad k > k_F^{N_\sigma} \\ E(N_\sigma, |\mathbf{k}| = k_F^{N_\sigma}) - E(N_\sigma - 1, \mathbf{k}) & \text{for} \quad k \leq k_F^{N_\sigma} \end{cases} \tag{6.4}$$

where the Fermi wavevector of a finite system $k_F^{N_\sigma}$ is different from the thermodynamic limit case

$$k_F^{N_\sigma} \neq k_F^{(N_\sigma)} = \left(6\pi^2 \frac{N_\sigma}{V}\right)^{1/3} \tag{6.5}$$

and it is understood that when changing N_σ we keep the population of the other spin component $\bar{\sigma}$ fixed (therefore we omitted the $N_{\bar{\sigma}}$ dependence in all QMC energies).
Since our code performs simulations normalized to $E_F^{(N_\uparrow)}$, let us rescale the dispersion relation by the same factor. Note that when $\sigma = \uparrow$ the total number of spin up fermion becomes $N_\uparrow \pm 1$, where $N_\uparrow$ is the number of spin up fermions in the ground state. So, when $\sigma = \uparrow$, adding or subtracting one fermion leads to the following normalization

$$\bar{E}(N_\uparrow \pm 1) = \frac{(N \pm 1) E_F^{(N_\uparrow \pm 1)}}{(N \pm 1) E_F^{(N_\uparrow \pm 1)}} \frac{E(N \pm 1)}{E_F^{(N_\uparrow)}} = (N \pm 1) E_{QMC}(N \pm 1) \left(\frac{N_\uparrow \pm 1}{N_\uparrow}\right)^{2/3} \tag{6.6}$$

while, when $\sigma = \downarrow$, one has

$$\bar{E}(N_\downarrow \pm 1) = \frac{(N \pm 1) E_F^{(N_\uparrow)}}{(N \pm 1) E_F^{(N_\uparrow)}} \frac{E(N \pm 1)}{E_F^{(N_\uparrow)}} = (N \pm 1) E_{QMC}(N \pm 1) \tag{6.7}$$

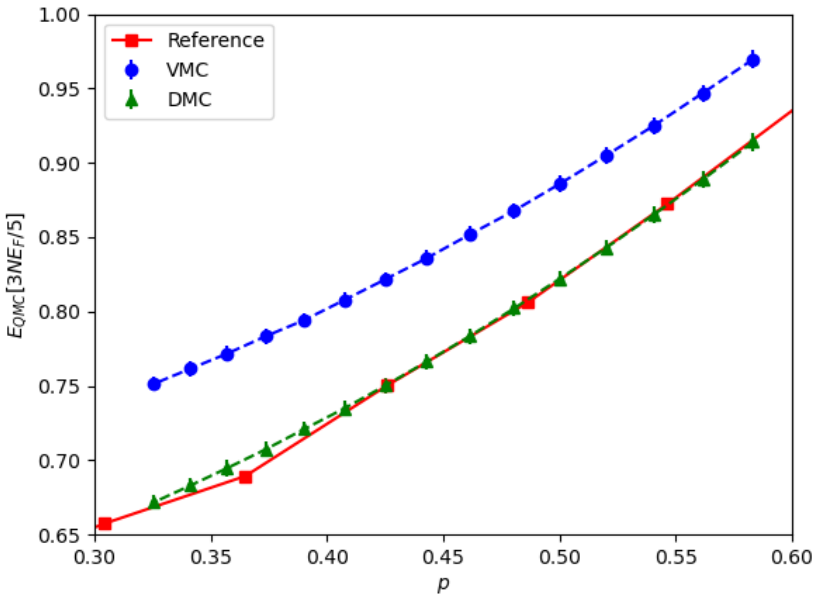


Figure 6.2: Quantum Monte Carlo energies for both variational (blue dots) and diffusion methods (green triangles), with also the results for DMC from [51] (red squares) labelled as reference.

where $\bar{E}$ is in Fermi energy units for $N_\uparrow$ in thermodynamic limit. So we have to rescale eq.(6.4) as

$$\bar{\varepsilon}_{N_\sigma}(\mathbf{k}) + \frac{\mu^{N_\sigma}}{E_F^{(N_\uparrow)}} = \begin{cases} \bar{E}(N_\sigma + 1, \mathbf{k}) - \bar{E}(N_\sigma, |\mathbf{k}| = k_F^{N_\sigma}) & \text{for} \quad k > k_F^{N_\sigma} \\ \bar{E}(N_\sigma, |\mathbf{k}| = k_F^{N_\sigma}) - \bar{E}(N_\sigma - 1, \mathbf{k}) & \text{for} \quad k \leq k_F^{N_\sigma} \end{cases} \tag{6.8}$$

6.2.2 Fermi liquid expansion

Let us now look for an expression that can be used to extract information from the QMC results. On the basis of Fermi liquid theory discussed in chapter 2.3, we can consider the dispersion relation for a finite system as

$$\varepsilon_{N_\sigma}(\mathbf{k}) \sim \frac{\hbar^2 k_F^{N_\sigma}}{m^*}\left(k - k_F^{N_\sigma}\right) = \frac{m}{m^*}\frac{\hbar^2 k_F^{N_\sigma}}{m}\left(k - k_F^{N_\sigma}\right) \tag{6.9}$$

where the wavevector $\mathbf{k}$ is assumed to be close to the Fermi sphere, with radius $k_F^{N_\sigma}$, which can in turn be expressed in terms of the quantum number $(n^{N_\sigma}_{\text{closed shell}})^2$ (which can be written as $n^2_{\text{closed shell}}$ omitting the dependence on N_σ) via

$$k_F^{N_\sigma} = \sqrt{n^2_{\text{closed shell}}}\frac{2\pi}{L} \tag{6.10}$$

while $\mathbf{k}$ is associated to some smaller or higher quantum number

$$n^2_{\text{shell}} > n^2_{\text{closed shell}}, \text{ or } n^2_{\text{shell}} < n^2_{\text{closed shell}} \ . \tag{6.11}$$

One can use this equation to fit our results, since we wish to extract information using the dispersion relation.
One can first try to fit the data with a first order function, however, one can also try to allow for the inclusion of points more distant from the Fermi sphere by incorporating a second order

correction. Note that this assumes there is not an inflection point at $k = k_F^{N_\sigma}$.
Since we use units of the Fermi energy of the spin-↑ ideal system, we are interested in the ratio between the Fermi energy of the ideal Fermi gas in the finite system $E_F^{N_\sigma}$ and the Fermi energy in the thermodynamic limit $E_F^{(N_\uparrow)}$. This ratio is

$$y_\sigma = \frac{E_F^{N_\sigma}}{E_F^{(N_\uparrow)}} = \frac{(k_F^{N_\sigma})^2}{\left(6\pi^2 \frac{N_\uparrow}{V}\right)^{2/3}} = \frac{n^2_{\text{closed shell}}}{N_\uparrow^{2/3}} \frac{4\pi^2}{(6\pi^2)^{2/3}} = n^2_{\text{closed shell}} \left(\frac{4\pi}{3N_\uparrow}\right)^{2/3} \tag{6.12}$$

Collecting $k_F^{N_\sigma}$, one can write the second order fit equation for the dimensionless dispersion $\bar{\varepsilon}(k)$ as

$$\bar{\varepsilon}(k) + \bar{\mu}^{N_\sigma} = \frac{2m}{m_\sigma^*} y_\sigma \left(\tilde{k} - 1\right) + \bar{c} \left(\tilde{k} - 1\right)^2 + \bar{d} \tag{6.13}$$

where

$$\tilde{k} = \frac{k}{k_F^{N_\sigma}} = \left[\frac{n^2_{\text{shell}}}{n^2_{\text{closed shell}}}\right]^{1/2} \tag{6.14}$$

and the fitting parameter $\bar{d}$ yields an estimate of the chemical potential in units of the Fermi energy of the spin-↑ ideal system.

6.3 Validation of the method

Let us consider here two checks of our method for extracting the effective masses.

6.3.1 Effective mass in the ideal Fermi gas

The first check of our method is for a non-interacting Fermi gas. In this case we know the exact results, and we verified only the effective mass of the spin-↑ components, in a system composed by $N_\uparrow = 57$ and $N_\downarrow = 27$. The expectation of effective mass $m^*_{\uparrow,\downarrow} = 1$ is satisfied as shown in figure 6.3 (within a statistical error of order 10^{-9}).
In this case the dispersion is a perfect parabola. Note that the fitting parameter $\bar{d} = 0.877$, which is fully consistent with the Fermi energy of the ideal gas in the finite system in units of the Fermi energy of the corresponding system in the thermodymanic-limit. Eq. (6.12) in this specific gas yields $5(4\pi/(3\cdot 57))^{2/3} \simeq 0.877$. Since we are satisfied by this result we can proceed with the Fermi polaron.

6.3.2 Effective mass in the Fermi polaron limit

In order to validate our method, at least in a limit of high polarization, we can perform a comparison with known results for the Fermi polaron effective mass. Let us concentrate on a system composed by $N_\uparrow = 57$ and $N_\downarrow = 1$ atoms, where we excite the impurity to different states increasing its quantum number n^2_{shell} from zero to eight. In this case we have $k_F^\downarrow = 0$ so our fit function reduces only to the second order term plus a constant, then as before we set units of $E_F^{(N_\uparrow)}$

$$\bar{\varepsilon}(k) + \bar{\mu}^\downarrow = \frac{\frac{\hbar^2}{2m^*_\downarrow}\left(\frac{2\pi}{L}\right)^2}{\frac{\hbar^2}{2m}\left(6\pi \frac{N^\uparrow}{V}\right)^{2/3}} \left(n^\downarrow_{\text{shell}}\right)^2 = \frac{m}{m^*_\downarrow} (n^\downarrow_{\text{shell}})^2 \left(\frac{4\pi}{3N_\uparrow}\right)^{2/3} + \bar{d} \tag{6.15}$$

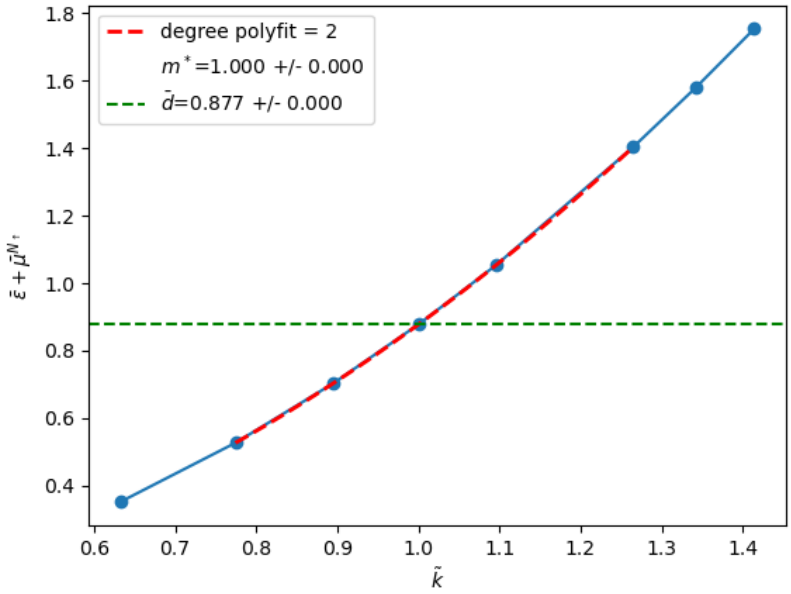


Figure 6.3: Dispersion relation for spin up quasiparticle for an ideal polarized Fermi gas composed by $N_\uparrow = 57$ and $N_\downarrow = 27$ atoms. Note that the fitting parameter $\bar{d} = 0.877$, which is fully consistent with the Fermi energy of the ideal gas in the finite system in units of the Fermi energy of the corresponding system in the thermodymanic-limit.

Here we are using the variable n^2_{shell} because in this way we can force the first order coefficient in k to be zero.
The datasets obtained by VMC and DMC methods contain both eight points, as shown in figure 6.4 for the second case. The dataset obtained by the first method yields $m^*_\downarrow/m = 0.999 \pm 0.001$, while the dataset from DMC simulations yields $m^*_\downarrow/m = 1.106 \pm 0.006$. Only the latter method is consistent with literature results. In particular, our result can be comparerd with the result by Pilati and Giorgini [61], who performed a Fixed node Monte-Carlo analysis of the Polarized Fermi Gas at Zero Temperature. They studied the equation of state for the system in different conditions of polarization, changing the number of $N_\downarrow$ atoms with fixed $N_\uparrow$, and extracted the effective mass $m^*/m = 1.09(2)$ from a functional relation for their results for the gas in high polarization limit.
As experimental result, one can see the work by Nascimbène et al. [96], where they found an effective mass of $m^*/m = 1.17(10)$. They determined the effective mass through the measurement of the oscillation frequency ω^* in a ultra-cold Fermi gas of ^{6}Li atoms at unitarity in a harmonic trapping potential. From this test in the polaron limit, one might be led to conclude that only DMC results could be trusted for the effective masses. We notice however that in the polaron problem only one spin-down fermion is present. As a consequence, there are no antisymmetrization requirements for spin-down fermions, and corresponding nodes in the wavefunction. For generic polarization, one has instead to take into account the antisymmetrization requirements, and the VMC and DMC methods will share the same nodal surfaces, due to the fixed node approximation. We expect that such a constraint may account for most of the effects on the effective masses that will be obtained by the two methods. These considerations motivate us to first perform a thorough VMC analysis of the effective masses as a function of the polarization, and consider the more computationally demanding DMC method only as a later stage.

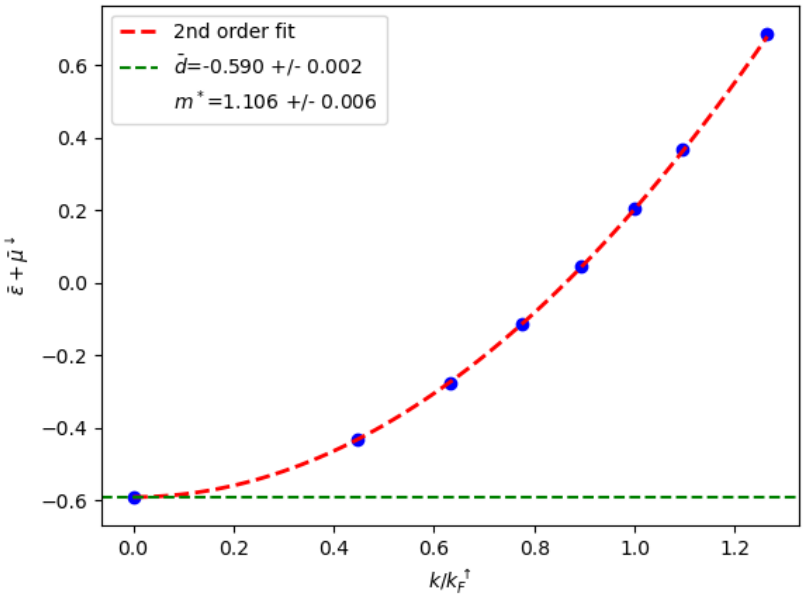


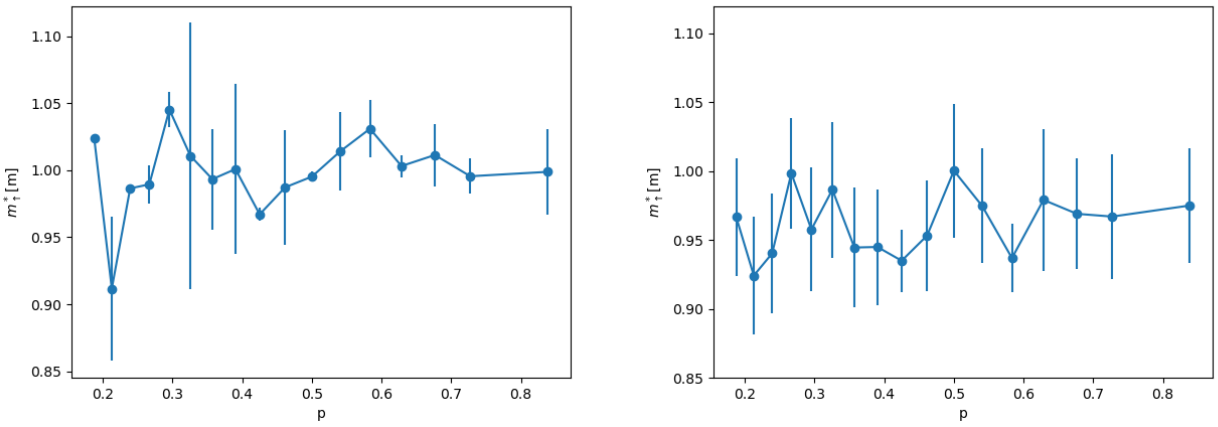
Figure 6.4: Dispersion relation for a single spin-down fermion in a system of $N_\uparrow = 57$ spin up fermions. The energies are obtained by DMC method.

6.4 Effective mass for different polarization

Let us now look at the behaviour of the effective masses for different polarizations p of the system. We start by considering the variational Monte Carlo method, then we will move to diffusion Monte Carlo. As we said before one can compute the fit of the dispersion relation with a polynomial of order one or two, this will have different effects, and will lead to different results, which are sometimes not compatible. Since we are interested in see some sort of breakdown of the Fermi liquid theory, we have to restrict our attention around k_F, . This requirement is, however, frustrated by the discrete nature of the wavevector $\mathbf{k}$ due to the maximum system sizes that are compatible with finite computational resources.
In order to reduce the finite size effects at least for one of the two spin species, we set the number of atoms for the species for which we compute the effective mass to corresponds to a closed shell in the reference simulation. Then we considered five simulations of the types requested in 6.2.1, two of them adding one fermion with $k > k_F^\sigma$ and two subtracting one with $k < k_F^\sigma$. The last one is the reference one, corresponding to the chosen closed shell.
We first considered the spin-↑ effective mass for system of $N_\uparrow = 57$ atoms in a closed shell. Changing the spin-↓ population we were able to inspect the discrete polarization range $[0.188, 0.839]$ ($N_\downarrow$ from 5 to 39). Then we considered the spin-down effective mass, so we set $N_\downarrow = 27$ changing now the spin-↑ component and inspecting the polarization range $[0.21, 0.64]$ ($N_\uparrow$ from 123 to 49).

6.4.1 Results for majority population

Let us start considering the spin up effective mass. If one operates on Monte Carlo energy results as explained in section 6.2.1, one can obtain the dispersion relation. Then one can analyse this dispersion performing one of the two fit described above.
When considering first a polynomial of degree one, we extracted the behaviour of the spin up



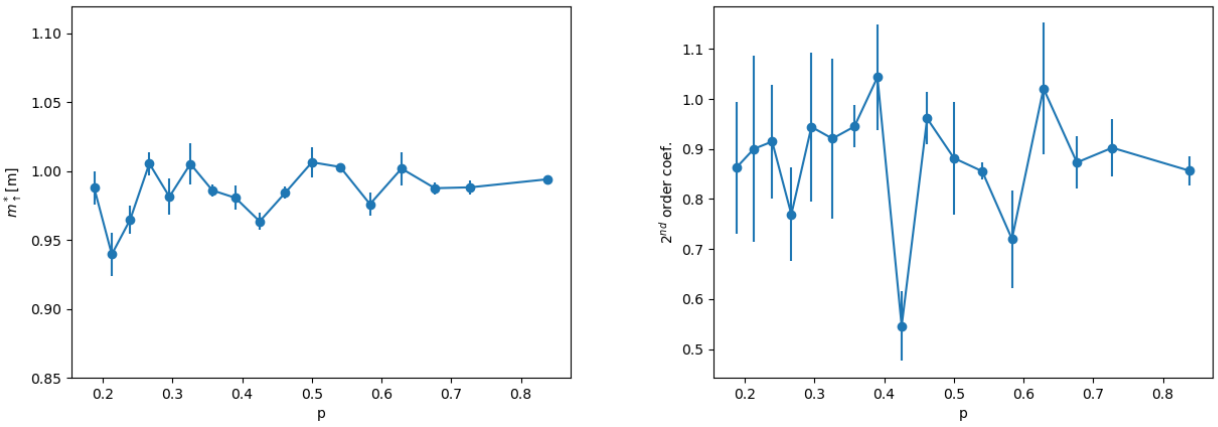
(a) (b)

Figure 6.5: Effective mass of spin-↑ quasi-particles in a system of $N_\uparrow = 57$ and $N_\downarrow$ variable for different number of points considered for fit: the case (a) is for three points while (b) is for five points.

effective mass $m^*_\uparrow$ reported in figure 6.5a-b, which differ for the number of points included in the fit, in first case three while in the second one five. One can see they are quite consistent and in the limit of large polarization they are compatible with our expectation. However, one can also consider the second degree polynomial, and use the eq.(6.13) to extract the effective mass from dispersion relation. The results are shown in figure 6.6a In this second case one can also inspect the behaviour of the second order coefficient in eq.(6.13) as function of polarization, trying to observe the evidence of the vanishing of this parameter which would indicate the presence of an inflection point at k^σ_F, as somewhat expected in the case of a flat dispersion around k^σ_F. As one can see in figure 6.6b, there is no strong evidence of this phenomena. Performing a t-student statistics test we obtain no significant compatibility with a second order coefficient (we can reject the hypothesis of non-significant coefficient within a confidence interval of 5%).
If one compares the goodness of fit for these two types of fit, one can see some differences between degree one (6.7a-b, for three points, and 6.7c-d, for degree two fits (fig. 6.7e-f, with five points). Generally speaking, we observe that the linear, five points fit is to be excluded since it fails to describe the dispersion curvature. The quadratic function with five points is able to fit better the data with respect to degree one: for all different numbers of spin-↓ atoms, these fits are good as one can see by χ^2/dof and relative p-value in figure 6.7e-f, but this could also lead to some issues. Since we are interested in the dispersion near the Fermi wavevector, considering a second order equation and including only few points around this value could result in overfitting the data, hiding then some important information. In addition the uncertainty of the parameters is lowered by this fitting procedure resulting in small error bars in figure 6.6a. In all cases there is no evidence of a divergence of the effective mass for spin up components.

6.4.2 Results for minority population

Let us now switch to the effective mass of the minority species, so we hold $N_\downarrow = 27$ and then, changing the number of spin-↑ components from $N_\uparrow = 123$ to 49, we can cover the polarization range $p \in [0.21, 0.64]$. Again one can analyse these dispersions fitting the data by polynomials of order one of two as above.



(a) (b)

Figure 6.6: a) Effective mass of spin-↑ component in a system of $N_\uparrow = 57$ with different polarizations p. b) Second order coefficient of fit equation as function of polarization.

Exactly as before, one can start by considering the polynomial of order one for fitting the dispersions. The resulting behaviour of the effective mass is reported in figure 6.8, depending on the number of points included for the fit, three points in panel (a) and five points in panel (b). One can see the results obtained by the two different fits are consistent with each other, and at large polarization both approach the VMC result in the polaronic limit.
Both fits also show an increase of the effective mass in the region of polarization around $p = 0.41$, followed by a decrease at lower polarization. Interestingly, this region is close to the expected position of the critical polarization ($p_c \simeq 0.435$) according to T-matrix calculation [1], even though clearly we cannot take it as an evidence of the predicted divergence.
Again one can also perform the fit with a polynomial of degree two, in this case one obtains the results in figure 6.9 As before the second option with more fitting parameters, not suprisingly, yield a better fit, especially in some regions of polarization, as one can see by comparison of their χ^2/dof and relative p-value in figure 6.10, where one can see fig. 6.10a-b for the degree one polynomial (three points), 6.10c-d (five points) and figures 6.10e-f for degree two with five points. For all types of fitting procedures the χ^2/dof is smaller at high polarization, hinting at a bending of the dispersion for intermediate polarizations. As we noted for the spin up case, we observe that linear fits with five points are to be excluded, since due to the finite size of the system, the spacing between the five points is large and the curvature of the dispersion is clearly important. This could be just the evidence of a too large freedom of this fitting in a limited range of points around k_F. As before the second order coefficient does not look compatible with zero.
However, this irregular behaviour of effective mass for both spin populations could be due to the presence of finite size effects still out of our control, and by enlarging the number of atoms of the system one would expect to obtain more regular results.

6.4.3 Larger systems

The systems considered previously were quite small, so the available k-points are significantly far from k_F, and these size effects may limit our possibilities of observing some non-Fermi liquid phenomena. So we considered larger systems, in order to reach bigger closed shells and reduce

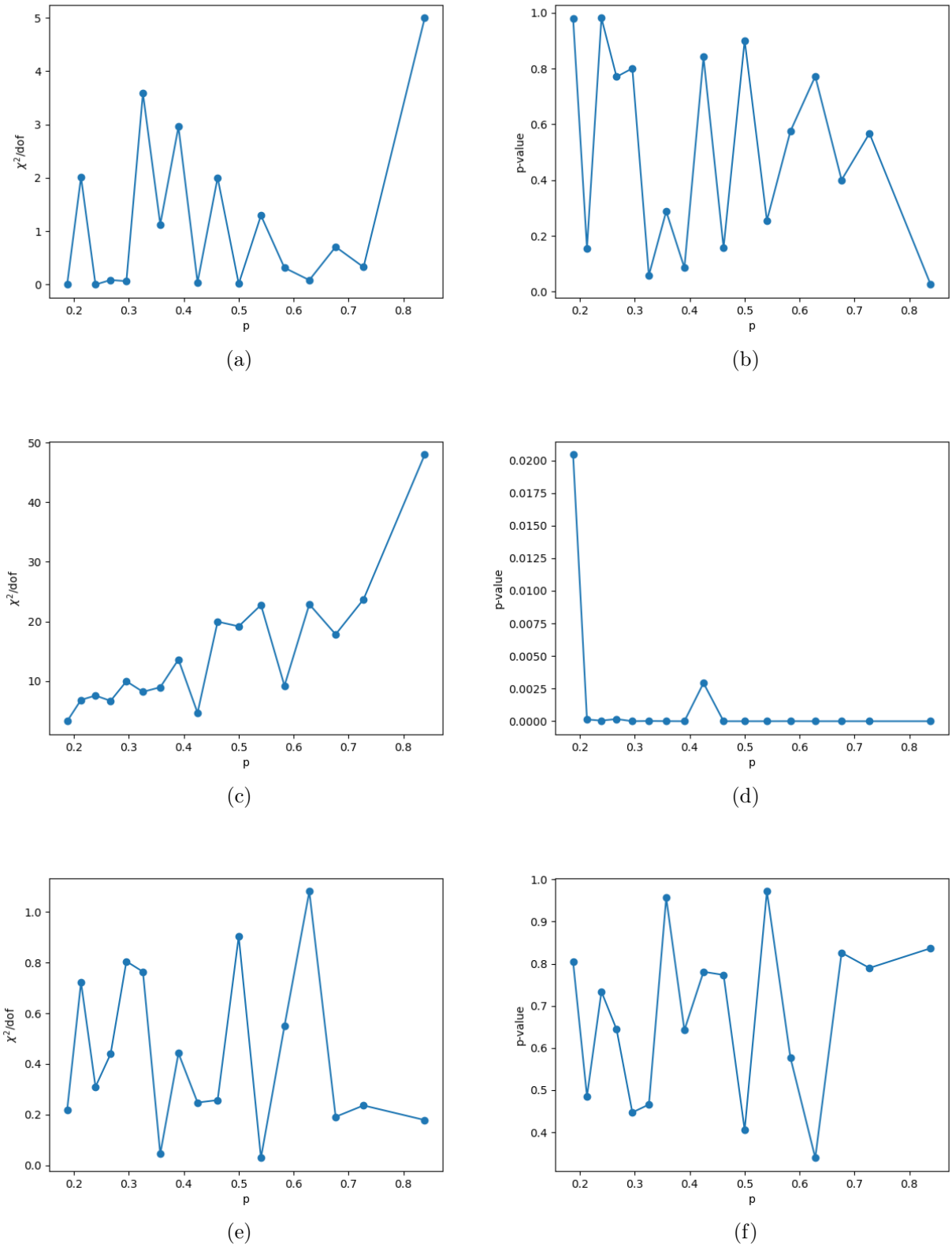
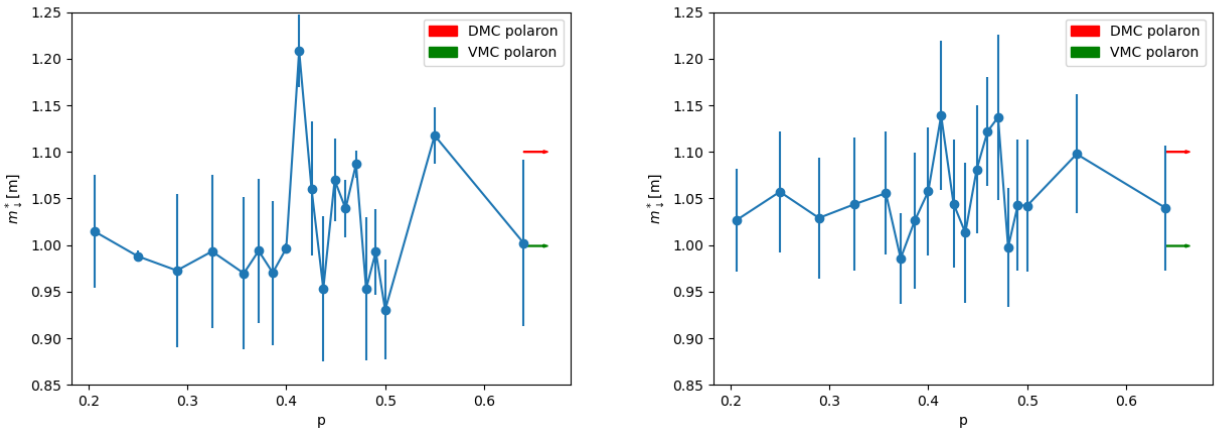
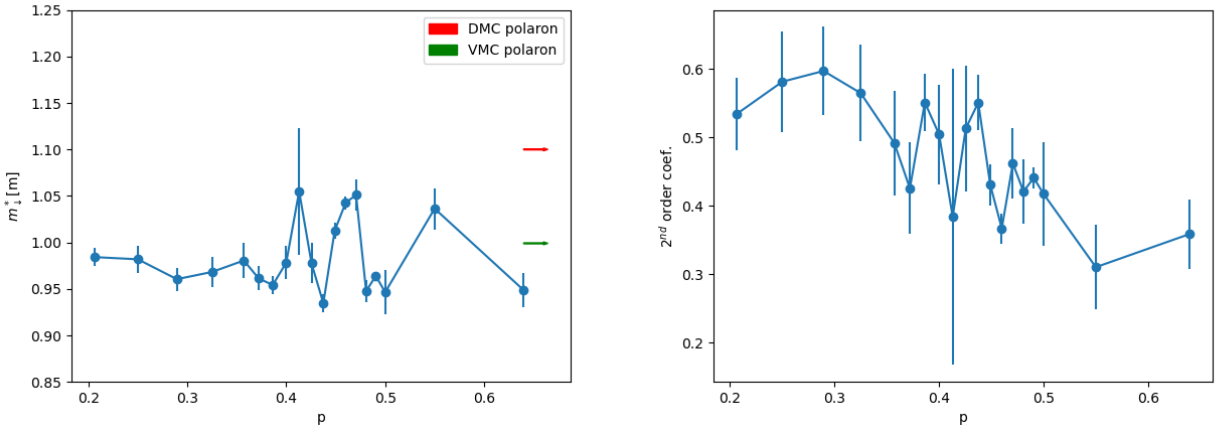


Figure 6.7: Resulting χ^2/dof and relative p-value from fit for different polarizations for spin up effective mass. a-b) Polynomial degree one and three points. c-d) Polynomial degree one and five points. e-f) Polynomial degree two and five points.



(a) (b)

Figure 6.8: Effective mass of spin-↓ quasi-particles in a system of $N_\downarrow = 27$ and $N_\uparrow$ variable for different number of points considered for fit: the case (a) is for three points while (b) is for five points.



(a) (b)

Figure 6.9: a) Effective mass of spin-↓ quasi-particles in a system of $N_\downarrow = 27$ and $N_\uparrow$ variable for different polarizations p for a quadratic fit. b) The second order coefficient of the quadratic fit.

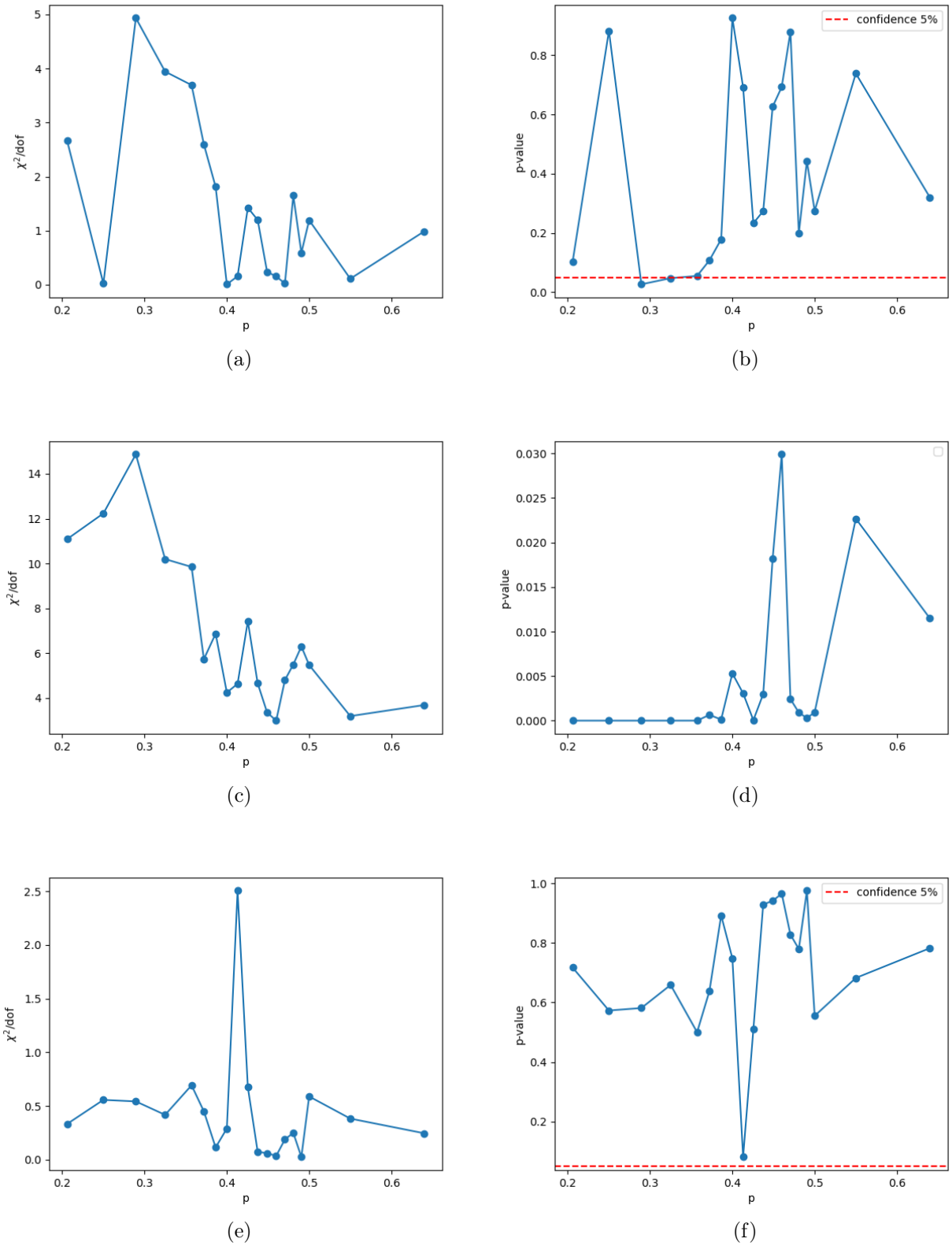
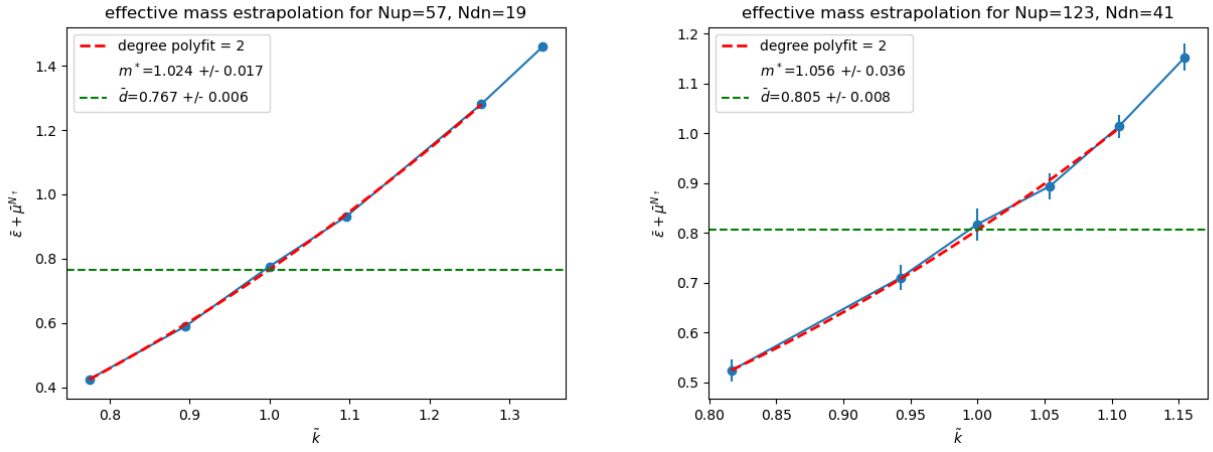


Figure 6.10: Resulting χ^2/dof and relative p-value from fit for different polarizations for spin down effective mass. a-b) Polynomial degree one and three points. c-d) Polynomial degree one and five points. e-f) Polynomial degree two and five points.



(a) (b)

Figure 6.11: Comparison of effective mass for a gas of atoms with same polarization but different number of components with the same number of iterations. a) Effective mass of spin-↑ component in a system of $N_\uparrow = 57$ and $N_\downarrow = 19$. b) Effective mass of spin-↑ component in a system of $N_\uparrow = 123$ and $N_\downarrow = 41$.

the relative distances of points in k-space. We expected improvements for both effective masses, but above all for the study of the effective mass of the minority population, where the initial value $N_\downarrow = 27$ may be too small, resulting in larger size effects.
We decided to move from $N_\uparrow = 57$ to $N_\uparrow = 123$ to repeat the analysis of spin up effective mass, while we moved from $N_\downarrow = 27$ to $N_\downarrow = 57$ for the spin down effective mass. These changes were somewhat forced. In fact, for the spin-up effective mass, since the first value of the quantum numbers n^2 available for the dispersion relation, that are bigger than $n^2 = 5$ (value for $N_\uparrow = 57$), are $n^2 = 6, 8, 9$, the jump from six to eight enforces the large increase of closed shell from $n^2 = 5$ to $n^2 = 9$. For the spin-down effective mass, even if the same limitation as before is still valid, we have also to face with the limited number of spin up atoms we can consider in our system to cover the polarization range, since a little increase in minority components reflects in a large increase of majority atoms for fixed polarization.
However, as one can see in figure 6.11, by comparing the dispersion relation of the new system for spin up effective mass, now the same level of error in QMC results lead to a much bigger uncertainty in energy. This affects strongly the fit results, since a small fluctuation of order 10^{-4} is sufficient to change the resulting effective mass. Then we decided to extend the simulation duration, increasing by four the number of steps in order to halve the errors. However this proved not enough for a reliable analysis of the effective mass. The same is true in the case of spin down effective mass. One can see in fig. 6.12 the results in this case for $N_\downarrow = 57$ and $N_\uparrow = 147$. What we expected from these dispersion was a flattening around the Fermi surface, with a corresponding divergence of the effective mass, since this is related to the derivative. Unfortunately even if it we were not able to observe this phenomena even if sometimes it looks like some flattening is occurring, the relative uncertainty was too big to exclude a random fluctuation and extract a reasonable effective mass from the fit. In fact, it may be necessary to go below an order of magnitude smaller error which means a number of iterations more than a hundred times greater.

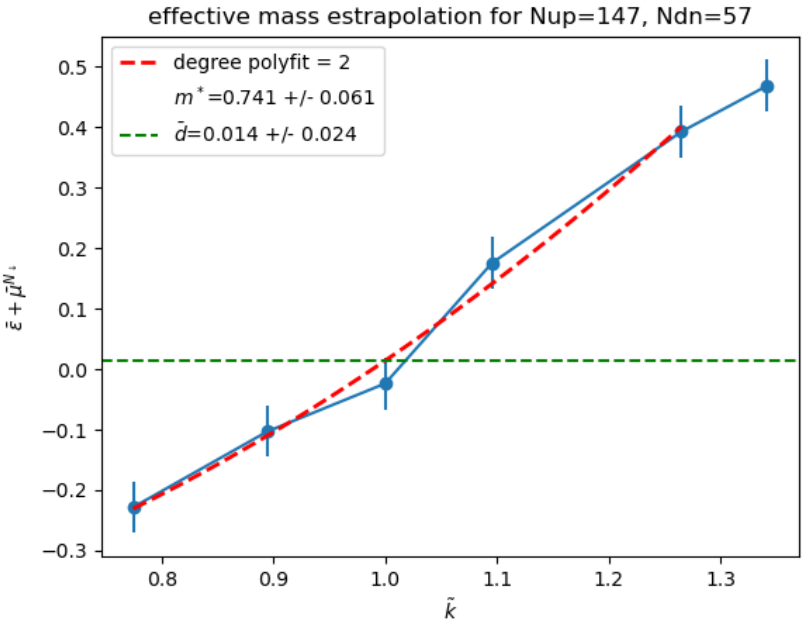
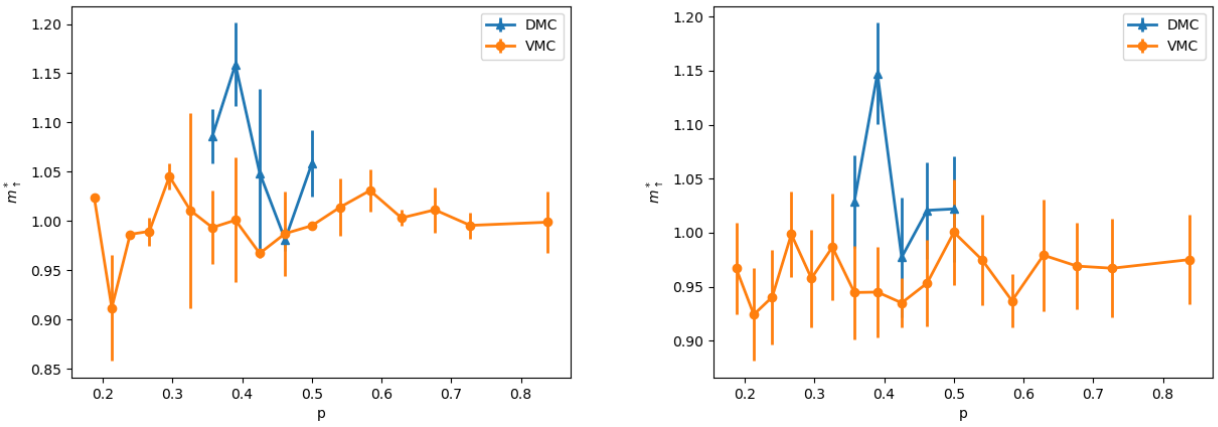


Figure 6.12: Dispersion relation for spin down quasiparticle in a system of $N_\uparrow = 147$ and $N_\downarrow = 57$.

6.4.4 Comparison between VMC and DMC

Even if VMC simulations are much faster than DMC, it took anyway quite a long time to produce and analyze the data presented so far for the effective mass. This precluded the possibility to inspect the same range of polarization with diffusion Monte Carlo. For this reason, performed a DMC analysis in a small polarization range covering the critical region determined by self-consistent t-matrix results in [1] and discussed in section 3.4.2. So, going back to the initial setup for spin-↑ effective mass, namely the system with $N_\uparrow = 57$, by changing $N_\downarrow$ from 19 to 27, one has $p \in [0.357, 0.5]$. We used six different time steps for a linear extrapolation to $dt = 0$, resulting in each case in a dispersion relation from which we extracted the effective mass as done before with VMC data. Then our results in this case are reported in figure 6.13 for fitting with a polynomial of degree one, and in fig6.14a for a polynomial of degree two, both in comparison with previous VMC results. From this comparison, we see that essentially for all fittings and polarizations here considered the effective masses obtained with the DMC method are slightly larger than those obtained by VMC, even though, for the fitting polynomial of order one, the differences between VMC and DMC results are within the error bars of the two methods for several polarizations.
It should be remarked that this small polarization range required more than a week of simulations on a 56 core machine. So an inspection over the whole range of polarization will require a simulation time of some months unless parallel simulation over multiple computational nodes is employed.



(a) (b)

Figure 6.13: Comparison between the effective mass behaviours as result of DMC simulations obtained from a fit function of order one for three (a) and five (b) points.

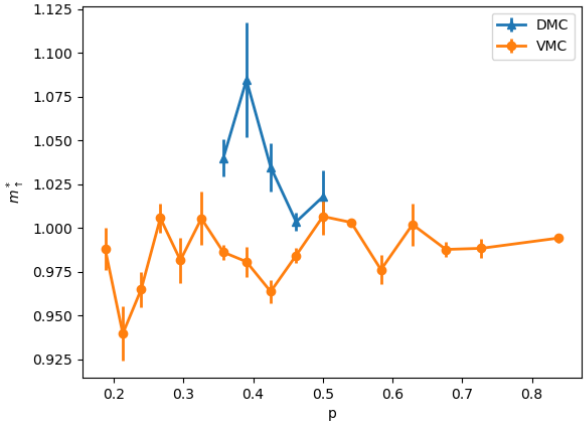


Figure 6.14: Spin up quasiparticle effective mass behaviour as function of polarization for fit equation of order two.

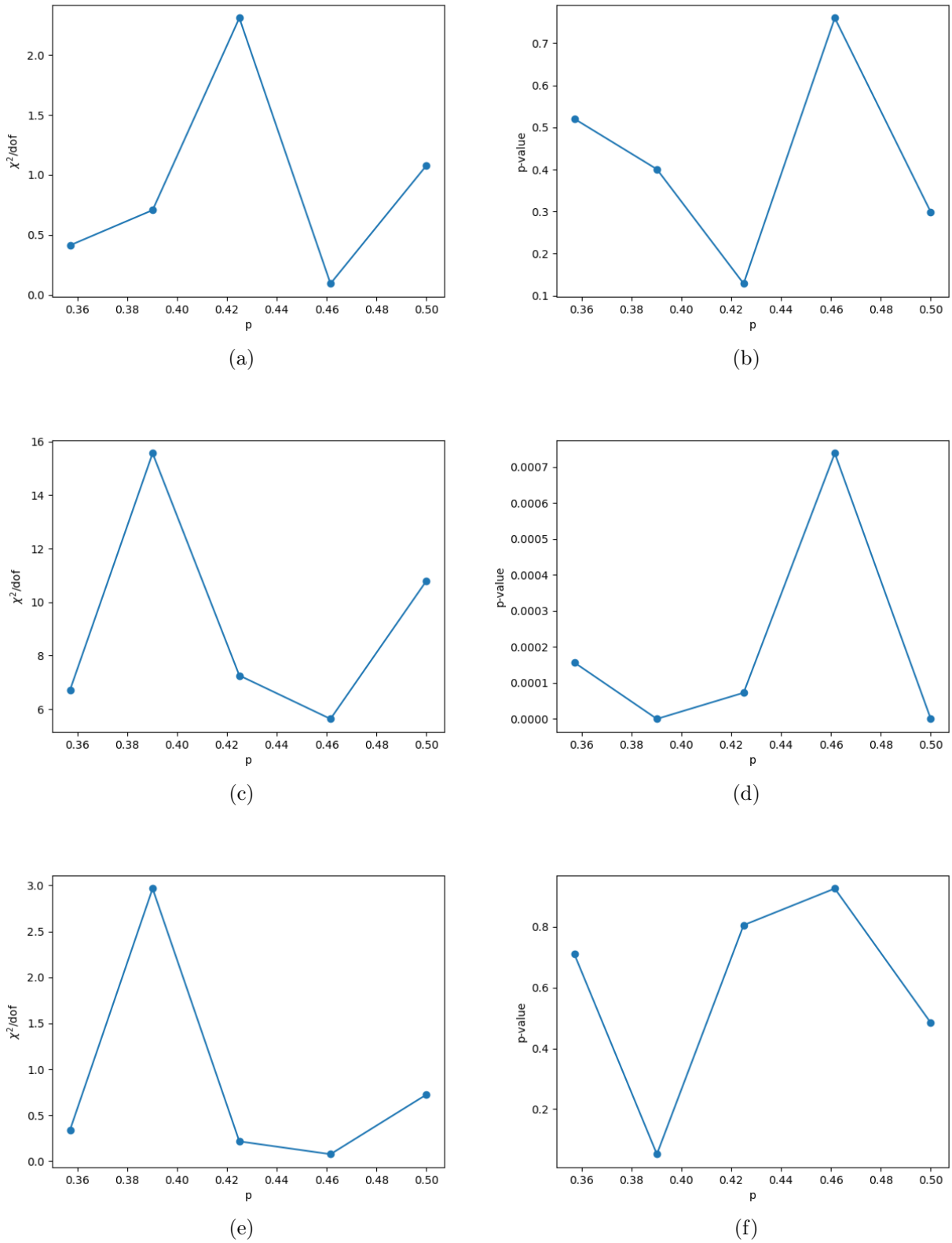


Figure 6.15: Resulting χ^2/dof and relative p-value from fit for different polarizations for spin up effective mass for DMC results. a-b) Polynomial degree one and three points. c-d) Polynomial degree one and five points. e-f) Polynomial degree two and five points.

CHAPTER 7

CONCLUSIONS AND PERSPECTIVES

In the previous chapter we presented the results obtained by simulating a polarized atomic Fermi gas at unitarity and at zero temperature. We applied Fermi liquid theory to study the effective mass for both spin components of the system. Our expectation was to highlight a critical behaviour of the gas as it transitions from high polarizations, where Fermi liquid theory is valid, to regions of low polarization, where this theory of weak excitations loses its meaning and validity. According to a recent theoretical work, this transition should be signaled by a sharp increase of the effective masses, and their divergence at a critical polarization. With our investigation we found only weak evidence of such a behaviour, and only for the effective mass of the minority component. We believe that the main issue, that prevented our simulations to provide a more convincing evidence, was the finite size of the systems we considered. The discretization in momentum space, stemming from the finite size of our simulations, together with the required precision for this analysis, prevented us to inspect the ideal configuration of the system, hiding the transition through a quantum critical point. In fact, it was not possible to extend the system size in order to check for a bending of the dispersion very close to the Fermi surface; this extension would have allowed for results less affected by finite-size effects. On the other hand, for larger systems, the energy differences shrank, making it impossible to achieve the same relative error threshold within the available computational costs. In fact, reducing the error to appropriate values would require a very large increase in computational time.

However, there are some possible future development that may lead to observe or definitively disproof, the prediction of diverging effective masses. A first obvious development could be to expand the size of the system in order to get closer to the Fermi momenta of the two spin species, but this would imply a very long simulation time. A better strategy would be overcoming size limitations by introducing twist-averaged boundary conditions as discussed in [97]. This could allow one to consider wavevectors $\mathbf{k}$ that are not determined by periodic boundary conditions, and which thus could be taken much closer to the Fermi surface.

A third development would be to perform more systematic and extensve DMC simulations, since we cannot exclude that moving from variational to diffusion Monte Carlo could significantly modify the results. An indication that some changes could indeed occur is given by the results obtained with DMC for the spin-up quasiparticle effective mass, albeit in a small polarization range and for a single size of the system.

Finally, a further future developments could be to improve the variational wave function. This upgrade would require an inclusion of backflow transformations, with the aim of acting espe-

cially on the nodal surface of the system. In this way, one could go beyond the Jastrow-Slater variational wave function with plane waves, that we implemented for our analysis of the effective mass.
In conclusion, in our work we have provided a thorough analysis of the calculation of the effective masses in a polarized atomic Fermi gas at unitarity, at zero temperature, with the VMC method with Jastrow-Slater wavefunctions for system sizes of the order of 100 particles. This analysis was complemented by a still preliminary investigation of the same problem with the DMC method.
We found only weak evidence of a predicted breakdown of Fermi liquid theory in the proximity of a quantum critical point. We have however highlighted the intrinsic limitations of the present study, while pointing out further improvements that will be necessary in future simulations to draw more definite conclusions on this problem.

CHAPTER 8

APPENDIX

8.1 Appendix A

We obtain here the relation between the magnitude of the energy variation and the feedback effect; in the following equations the momentum dependence will be omitted whenever not necessary. Let us start by considering the effect of an external potential (or field) that leads to the presence of a new term in quasi-particle energy variation, this term is $\delta\hat{\varepsilon}^{(0)}$ and must be added to the energy variation generated by interactions, resulting in a total quasi-particle energy variation

$$\delta\hat{\varepsilon} = \delta\hat{\varepsilon}^{(0)} + \delta\hat{\xi} = \delta\hat{\varepsilon}^{(0)} + \mathrm{Tr}\int \hat{f}(\mathbf{p},\mathbf{p}')\delta\hat{n}(\mathbf{p}')\frac{d\mathbf{p}'}{(2\pi)^3} \tag{8.1}$$

that reflects on quasiparticle occupancies as

$$\hat{n} = f(\hat{\varepsilon}^{(0)} + \delta\hat{\varepsilon}) \approx f(\hat{\varepsilon}^{(0)}) + f'(\hat{\varepsilon}^{(0)})\delta\hat{\varepsilon} \tag{8.2}$$

where here f indicate the Fermi distribution. As the temperature is lowered to zero f tends to be an Heaviside function, and its derivative a delta function

$$\hat{n} = \theta(\hat{\varepsilon}^{(0)}) + \left[-\delta(\hat{\varepsilon}^{(0)})\delta\hat{\varepsilon}\right] \tag{8.3}$$

where $\delta\hat{n} = -\delta(\hat{\varepsilon}^{(0)})\delta\hat{\varepsilon}$ represents the polarization of the Fermi surface. This effect goes inside the variation $\delta\hat{\varepsilon}$, resulting in a self-consistent equation

$$\delta\hat{\varepsilon} = \delta\hat{\varepsilon}^{(0)} - \mathrm{Tr}\int \hat{f}(\mathbf{p},\mathbf{p}')\delta(\hat{\varepsilon}^{(0)}(\mathbf{p}'))\delta\hat{\varepsilon}(\mathbf{p}')\frac{d\mathbf{p}'}{(2\pi)^3} \tag{8.4}$$

The feedback process preserves the symmetry of the external perturbation, but change the intensity by a relative Landau parameter, that depends on the perturbation. Isotropic effects depend on $l = 0$ coefficients, while for different symmetries $l > 0$.
In order to obtain a relation between the magnitude of the energy variation and the feedback effect, we can consider a quasiparticle potential that has a particular multipole symmetry, then the bare change of quasiparticle energy can be written as

$$\delta\hat{\varepsilon}^{(0)}(\mathbf{p}) = v_l Y_{lm}(\hat{\mathbf{p}}) \tag{8.5}$$

where Y_{lm} is a spherical harmonic. This change leads to a total variation of energy $\delta\hat{\varepsilon}$ that has the same symmetry, while the magnitude is changed from v_l to t_l,

$$\delta\hat{\varepsilon}(\mathbf{p}) = t_l Y_{lm}(\hat{\mathbf{p}}) \tag{8.6}$$

which leads to the variation of the occupation number

$$\delta\hat{n} = -t_l Y_{lm}(\hat{\mathbf{p}})\delta(\hat{\varepsilon}^{(0)}) \tag{8.7}$$

that affects the energy variation generated by interactions as

$$\delta\hat{\xi} = -\mathrm{Tr}\int \hat{f}(\mathbf{p},\mathbf{p}')\left(t_l Y_{lm}(\hat{\mathbf{p}}')\delta(\hat{\varepsilon}^{(0)}(\mathbf{p}')\right)\frac{d\mathbf{p}'}{(2\pi)^3} \tag{8.8}$$

Now, if one considers a system where the trace of the spin dependence generates just a multiplicative constant, and if one decomposes the interaction function into its symmetrical and anti-symmetrical components, the trace operator removes the anti-symmetrical component and the energy variation becomes

$$\delta\xi(\mathbf{p}) = -t_l s\int f^s(\mathbf{p},\mathbf{p}')Y_{lm}(\hat{\mathbf{p}}')\delta(\varepsilon^{(0)}(\mathbf{p}'))\frac{d\mathbf{p}'}{(2\pi)^3} \tag{8.9}$$

where s is a factor that depends on the spin degeneracy: if we consider the case of a spin-1/2 system, $s = 2$. It is useful to transform the integral over $\mathbf{p}'$ into an angular average over the Fermi surface, this can be done by considering that $\delta(\varepsilon^{(0)}(p))$ acts as a projector onto the fermi surface, so the transformation of the integrals is

$$2\int \delta(\varepsilon^{(0)}(\mathbf{p}))\frac{d\mathbf{p}}{(2\pi)^3} \to \frac{m^* k_F}{\pi^2}\int \frac{d\Omega_{\hat{\mathbf{p}}}}{4\pi} \tag{8.10}$$

Then, by applying this transformation to equation 8.9, we obtain

$$\delta\xi(\mathbf{p}) = -t_l\frac{m^* k_F}{\pi^2}\int \frac{d\Omega_{\mathbf{p}'}}{4\pi}f^s(\mathbf{p},\mathbf{p}')Y_{lm}(\hat{\mathbf{p}}') = -t_l\int \frac{d\Omega_{\mathbf{p}'}}{4\pi}F^s(\mathbf{p},\mathbf{p}')Y_{lm}(\hat{\mathbf{p}}') \tag{8.11}$$

If we are in a rotational symmetric system, F depends only on the angle between momenta $\mathbf{p}$ and $\mathbf{p}'$, and we can expand in Legendre polynomials this function, that can be further decomposed in spherical harmonics

$$F^s(\cos\theta(\mathbf{p},\mathbf{p}')) = \sum_{l=0}^{\infty} F_l^s P_L\left(\cos\theta(\mathbf{p},\mathbf{p}')\right) = \frac{4\pi}{2l+1}\sum_{lm}^{\infty} F_l^s Y_{lm}(\hat{\mathbf{p}})Y_{lm}^*(\hat{\mathbf{p}}') \tag{8.12}$$

Then, by means of orthogonality of the spherical harmonics, eq.(8.11) turns into the following form

$$\begin{aligned}\delta\xi(\mathbf{p}) &= -t_l\frac{4\pi}{2l+1}\sum_{l'm'}^{\infty} F_{l'}^s Y_{l'm'}(\hat{\mathbf{p}})\int \frac{d\Omega_{\mathbf{p}'}}{4\pi}Y_{l'm'}^*(\hat{\mathbf{p}}')Y_{lm}(\hat{\mathbf{p}}') \\ &= -\frac{t_l}{2l+1}\sum_{l'm'}^{\infty} F_{l'}^s Y_{l'm'}(\hat{\mathbf{p}})\int d\Omega_{\mathbf{p}'}\delta_{ll'}\delta_{mm'} = -\frac{t_l}{2l+1}F_l^s Y_{lm}(\hat{\mathbf{p}})\end{aligned} \tag{8.13}$$

Finally we obtain the relation

$$\delta\varepsilon(\mathbf{p}) = v_l Y_{lm}(\hat{\mathbf{p}}) - \frac{t_l}{2l+1}F_l^s Y_{lm}(\hat{\mathbf{p}}) = t_l Y_{lm}(\hat{\mathbf{p}}) \tag{8.14}$$

which allows one to obtain

$$t_l = \frac{v_l}{1+\frac{F_l}{2l+1}} \tag{8.15}$$

8.2 Appendix B

We will discuss here the proof of the relation between mass and effective mass of quasiparticles by following the original work by Landau [13].
One has to set the equivalence of momentum for unit volume of the liquid, and the density of its mass flow. Because the velocity of a quasiparticle is $\mathbf{v} = \partial\varepsilon/\partial\mathbf{p}$, the flow can be computed as

$$\mathrm{Tr}\int d\boldsymbol{\tau}\hat{n}\frac{\partial\varepsilon}{\partial\mathbf{p}} \tag{8.16}$$
$$d\boldsymbol{\tau} = \frac{d\mathbf{p}}{(2\pi)^3}$$

Since the number of quasiparticles with momenta close to the Fermi momentum is the same as that of real particles, the mass transport can be computed as the current of quasiparticles number multiplied by the bare mass. So the equation to solve is

$$\mathrm{Tr}\int d\boldsymbol{\tau}\mathbf{p}\hat{n} = m\mathrm{Tr}\int d\boldsymbol{\tau}\frac{\partial\hat{\varepsilon}}{\partial\mathbf{p}}\hat{n} \tag{8.17}$$

Let us make the assumption that the spin dependence leads just to a multiplicative factor after the trace, then let us consider the variation of both side of previous eq.(8.17), by introducing the previous results for energy and interactions function. So on the right-hand side of eq.(8.17) we have

$$m\int d\boldsymbol{\tau}\frac{\partial\varepsilon(\mathbf{p})}{\partial\mathbf{p}}\delta n(\mathbf{p}) + m\int d\boldsymbol{\tau}n(\mathbf{p})\int d\boldsymbol{\tau}'\frac{\partial f(\mathbf{p},\mathbf{p}')}{\partial\mathbf{p}}\delta n(\mathbf{p}') = \\ m\int d\boldsymbol{\tau}\frac{\partial\varepsilon(\mathbf{p})}{\partial\mathbf{p}}\delta n(\mathbf{p}) - m\int d\boldsymbol{\tau}\delta n(\mathbf{p})\int d\boldsymbol{\tau}' f(\mathbf{p},\mathbf{p}')\frac{\partial n(\mathbf{p}')}{\partial\mathbf{p}'} \tag{8.18}$$

where in the last line a change of variables $(\mathbf{p},\mathbf{p}')$ and an integration by parts were performed. Due to the arbitrariness of δn, we can extract

$$\mathbf{p} = m\frac{\partial\varepsilon}{\partial\mathbf{p}} - m\int d\boldsymbol{\tau}' f(\mathbf{p},\mathbf{p}')\frac{\partial n(\mathbf{p}')}{\partial\mathbf{p}'} \tag{8.19}$$

Let us consider again the case of momenta restricted in a narrow range around k_F, so we can simplify the distribution as a step function $n(\mathbf{p}') = \theta(\mathbf{p}')$. Then the derivative of $n(\mathbf{p}')$ becomes

$$\frac{\partial n(\mathbf{p}')}{\partial\mathbf{p}'} = \frac{\partial\theta(\mathbf{p}')}{\partial\mathbf{p}'} = -\frac{\mathbf{p}'}{|\mathbf{p}'|}\delta(p'-k_F) \tag{8.20}$$

Because of the restricted range of the momenta, we can approximate the energy derivative as the quasiparticle velocity on the Fermi surface $v_F = k_F/m^*$. So finally

$$\frac{\mathbf{p}}{m} = \frac{\mathbf{k}_F}{m^*} + \int d\boldsymbol{\tau}' f(\mathbf{p},\mathbf{p}')\frac{\mathbf{p}'}{|\mathbf{p}'|}\delta(p'-k_F) \tag{8.21}$$

Now let us multiply both sides by $\mathbf{p}$, and replace the function $\delta(p'-k_F)$ with $v_F\delta(\varepsilon(\mathbf{p}))$. Then we can remove the angular dependence and set $p = p' = k_F$, after that we can replace the volume integral with a solid angle integral as discussed in appendix A

$$\frac{\mathbf{p}^2}{m} = \frac{\mathbf{p}\cdot\mathbf{k}_F}{m^*} + \frac{k_F^2}{m^*}\int d\boldsymbol{\tau}' f(\mathbf{p},\mathbf{p}')\frac{\mathbf{p}\cdot\mathbf{p}'}{k_F|\mathbf{p}'|}\delta(\varepsilon(\mathbf{p})) = \frac{k_F^2}{m^*} + \frac{k_F^2}{m^*}\frac{m^*k_F}{\pi^2}\int\frac{d\Omega_{\mathbf{p}'}}{4\pi}f(\mathbf{p},\mathbf{p}')\cos\theta \tag{8.22}$$

Then the final equation is

$$\frac{1}{m} = \frac{1}{m^*} + \frac{1}{m^*} \int \frac{d\Omega_{\mathbf{p}'}}{4\pi} F(\cos\theta)\cos\theta \tag{8.23}$$

This relation connects the mass of the element of the liquid with the mass of the quasi-particle. If we expand the interaction function in Legendre polynomials, and keeping only the $l = 1$ spin symmetric interaction coefficient, we can obtain

$$\frac{1}{m} = \frac{1}{m^*} + \frac{F_1^s}{3m^*} \tag{8.24}$$

8.3 Appendix C

In this appendix we will discuss the validity of the algorithm by Metropolis et al. proved by Rosenbluth [78] focusing on the conservation of probability. In fact, he proved that for a classical system, under the conditions of ergodicity and detailed balance, probability flows through configuration space in such a way that the average squared-deviation of the probability density in canonical ensemble not only becomes zero but it does so monotonically.
Let us start by considering a set of M points that will be moved simultaneously through configuration space. Let us define the position vector in configuration space $\mathbf{r}$ with a density $\rho(\mathbf{r})$. Every time a transition occurs in our process, if we keep fixed the total number of elements, those points N_i that enter one specific region of space $d\mathbf{r}$ must come from any other region $d\mathbf{s}$. As consequence after another step some points N_o will leave $d\mathbf{r}$. Then the total variation of points in region $d\mathbf{r}$ after one step in our Markov chain (from one configuration we will label as 1 to one second labelled as 2) is

$$\rho_2(\mathbf{r})d\mathbf{r} - \rho_1(\mathbf{r})d\mathbf{r} = N_i - N_o \tag{8.25}$$

Let us define $P_{sr}d\mathbf{r}$ as the probability to move from any region $\mathbf{s}$ into $d\mathbf{r}$, and since it is also possible that some elements stay fixed in $\mathbf{r}$, we have

$$P_{rr} + \int P_{rs}d\mathbf{r} = 1 \tag{8.26}$$

where P_{rr} is the probability to remain in $\mathbf{r}$, that could be finite. The previous numbers of points are

$$\begin{aligned} N_i &= d\mathbf{r} \int d\mathbf{s}\rho_1(\mathbf{s})P_{sr} \\ N_o &= \rho_1(\mathbf{r})d\mathbf{r} \int d\mathbf{s}P_{rs} \end{aligned} \tag{8.27}$$

So eq.(8.25) can be written as

$$\rho_2(\mathbf{r}) - \rho_1(\mathbf{r}) = \int d\mathbf{s}\left[\rho_1(\mathbf{s})P_{sr} - \rho_1(\mathbf{r})P_{rs}\right] \tag{8.28}$$

P_{rs} must fulfil two prescriptions in order to be considered a valid move, the first is the ergodicity and the second is the detailed balance

$$P_{rs}w(\mathbf{r}) = P_{sr}w(\mathbf{s}) \leftrightarrow P'_{rs} = P'_{rs} \tag{8.29}$$

where we defined the probability of occupation of $\mathbf{r}$ as $w(\mathbf{r})$. Let us set also

$$\rho'(\mathbf{r}) = \rho(\mathbf{r})w^{-1}(\mathbf{r}) \tag{8.30}$$

Now we can rewrite eq.(8.28) as

$$\rho_2(\mathbf{r}) - \rho_1(\mathbf{r}) = \int d\mathbf{s} P'_{rs} \left[\rho'_1(\mathbf{s}) - \rho'_1(\mathbf{r})\right] \tag{8.31}$$

If the distribution $\rho(\mathbf{r})$ is the canonical distribution then the spatial dependence is lost. Let us consider now the quantity

$$X = \int d\mathbf{r}\rho^2(\mathbf{r})w^{-1}(\mathbf{r}) = \int d\mathbf{r}\rho'^2(\mathbf{r})w(\mathbf{r}) \tag{8.32}$$

One can subtract a constant from X

$$X = \int d\mathbf{r} w(\mathbf{r}) \left[\rho'^2(\mathbf{r}) - \rho'^2_0\right] \tag{8.33}$$

where ρ'_0 is the value assumed by $\rho'(\mathbf{r})$ for the equilibrium distribution. Because of the conservation of probability

$$\int \rho d\mathbf{r} = \int \rho_0 d\mathbf{r} \tag{8.34}$$

their difference is the same as the difference of the square or the squared difference, then

$$X = \int d\mathbf{r} w(\mathbf{r}) \left[\rho'(\mathbf{r}) - \rho'_0\right]^2 \tag{8.35}$$

Now X represents an average squared deviation of the distribution function from the canonical distribution. Now we can show that any move that obeys the above restrictions will either decrease X, or leave it unchanged, but it is left unchanged only if $\rho'(\mathbf{r}) =$ constant, i.e if we have the equilibrium distribution. Let us consider the variation of X for two successive moves

$$\begin{aligned}\Delta X &= \int d\mathbf{r} w(\mathbf{r}) \left[\rho'^2_2(\mathbf{r}) - \rho'^2_1(\mathbf{r})\right] = \int d\mathbf{r} w(\mathbf{r}) \left[\rho'_2(\mathbf{r}) - \rho'_1(\mathbf{r})\right] \left[2\rho'_1(\mathbf{r}) + (\rho'_2(\mathbf{r}) - \rho'_1(\mathbf{r}))\right] \\ &= 2\int d\mathbf{r} w(\mathbf{r})\rho'_1(\mathbf{r}) \left[\rho'_2(\mathbf{r}) - \rho'_1(\mathbf{r})\right] + \int d\mathbf{r} w(\mathbf{r}) \left[\rho'_2(\mathbf{r}) - \rho'_1(\mathbf{r})\right]^2 = I + II\end{aligned} \tag{8.36}$$

Let us substitute now eq.(8.31) and (8.30) in first integral, and eq.(8.29) and (8.31) in second one, to obtain

$$I = 2\int d\mathbf{r}\rho'_1(\mathbf{r}) \left[\rho_2(\mathbf{r}) - \rho_1(\mathbf{r})\right] = 2\iint d\mathbf{r}d\mathbf{s}\rho'_1(\mathbf{r})P'_{rs} \left[\rho'_1(\mathbf{s}) - \rho'_1(\mathbf{r})\right] \tag{8.37}$$

Because of integration in both $\mathbf{r}$ and $\mathbf{s}$, one can interchange them

$$I = 2\iint d\mathbf{r}d\mathbf{s}\rho'_1(\mathbf{s})P'_{sr} \left[\rho'_1(\mathbf{r}) - \rho'_1(\mathbf{s})\right] \tag{8.38}$$

where $P'_{rs} = P'_{sr}$, we can take the average of the two expressions

$$\begin{aligned}I &= \frac{1}{2}\iint d\mathbf{r}d\mathbf{s}2P'_{rs} \left[\rho'_1(\mathbf{r}) \left(\rho'_1(\mathbf{s}) - \rho'_1(\mathbf{r})\right) + \rho'_1(\mathbf{s}) \left(\rho'_1(\mathbf{r}) - \rho'_1(\mathbf{s})\right)\right] \\ &= \iint d\mathbf{r}d\mathbf{s}P'_{rs} \left[\rho'_1(\mathbf{r})\rho'_1(\mathbf{s}) - \rho'^2_1(\mathbf{r}) + \rho'_1(\mathbf{s})\rho'_1(\mathbf{r}) - \rho'^2_1(\mathbf{s})\right] \\ &= -\iint d\mathbf{r}d\mathbf{s}P'_{rs} \left[\rho'_1(\mathbf{r}) - \rho'_1(\mathbf{s})\right]^2 = -\iint d\mathbf{r}d\mathbf{s}w(\mathbf{r})P_{rs} \left[\rho'_1(\mathbf{r}) - \rho'_1(\mathbf{s})\right]^2\end{aligned} \tag{8.39}$$

substitute now eq.(8.29) and (8.31) in second integral to obtain

$$II = \int d\mathbf{r} w(\mathbf{r}) \left[\rho_2'(\mathbf{r}) - \rho_1'(\mathbf{r})\right]^2 = \int d\mathbf{r} w^{-1}(\mathbf{r}) \left[\rho_2(\mathbf{r}) - \rho_1(\mathbf{r})\right]^2 =$$
$$\int d\mathbf{r} w^{-1}(\mathbf{r}) \left[\int ds P'_{rs} \left(\rho_1'(\mathbf{s}) - \rho_1'(\mathbf{r})\right)\right]^2 = \int d\mathbf{r} w(\mathbf{r}) \left[\int d\mathbf{s} P_{rs} \left(\rho_1'(\mathbf{s}) - \rho_1'(\mathbf{r})\right)\right]^2 \quad (8.40)$$

So finally

$$\Delta X = I + II = -\iint d\mathbf{r} d\mathbf{s} w(\mathbf{r}) P_{rs} \left[\rho_1'(\mathbf{r}) - \rho_1'(\mathbf{s})\right]^2 + \int d\mathbf{r} w(\mathbf{r}) \left[\int d\mathbf{s} P_{rs} \left(\rho_1'(\mathbf{s}) - \rho_1'(\mathbf{r})\right)\right]^2$$
$$= \int d\mathbf{r} w(\mathbf{r}) \left(\left[\int d\mathbf{s} P_{rs} \left(\rho_1'(\mathbf{s}) - \rho_1'(\mathbf{r})\right)\right]^2 - \int d\mathbf{s} P_{rs} \left[\rho_1'(\mathbf{r}) - \rho_1'(\mathbf{s})\right]^2 \right) \quad (8.41)$$

where, because of the square in the integral, we can switch the order of $\rho(\mathbf{r})$ and $\rho(\mathbf{s})$.
If we now define, for a given point $\mathbf{r}$,

$$P_{rs} d\mathbf{s} = dy \quad (8.42)$$

$$\int P_{rs} d\mathbf{s} = y \quad (8.43)$$

the previous equation becomes

$$\Delta X = \int d\mathbf{r} w(\mathbf{r}) \left\{ \overline{\left(\rho_1'(\mathbf{r}) - \rho_1'(y)\right)}^2 y^2 - \overline{\left(\rho_1'(\mathbf{r}) - \rho_1'(y)\right)^2} y \right\} \quad (8.44)$$

where the bar denotes an average over y and the curly brackets is a function of $\mathbf{r}$. Since the average square of a quantity is always greater than, or equal to, the squared average

$$\overline{\left(\rho_1'(\mathbf{r}) - \rho_1'(y)\right)}^2 \leq \overline{\left(\rho_1'(\mathbf{r}) - \rho_1'(y)\right)^2} \quad (8.45)$$

and from eq.(8.26) we have that $y \leq 1$, and so $y^2 \leq y$. Thus for all $\mathbf{r}$ the curly brackets is negative or zero and then $\Delta X \leq 0$. The equality is valid when $\rho_1'(\mathbf{r}) = \rho_1'(\mathbf{s})$ for all pairs of points $\mathbf{r}$ and $\mathbf{s}$, i.e, if we have the equilibrium distribution. Thus a move will always decrease X unless we have the equilibrium distribution. Hence after a sufficient large number of steps, X must come arbitrarily close to its minimum value, i.e., the density $\rho(\mathbf{r})$ will come arbitrarily close to the equilibrium distribution.
Previously we considered an ensemble of points, but we may be interested in only one of them. So for any initial position $\mathbf{r}_0$ in configuration space, after a sufficiently large number of steps, the probability that the point is anywhere in configuration space is simply the equilibrium probability, irrespective of the initial position. If now we call N the sufficient number of moves for a system to thermalize, we could get a correct canonical average if we consider the sampling of the points in position N-th, $2N$-th, ... However it is equally correct any positive shift of that position, like $N+1$, $N+2$, ... and still correct is to average all these averages, so taking the average of all positions after N initial moves. With the prescription that each move obeys the two initial restrictions.
This observation completes the proof of the theorem that an ergodic Markov process, with a detailed balance condition, must converge to the target distribution. It also follows that the eigenvalue with the largest magnitude is one and that the corresponding eigenvector is unique and equal to the correct distribution, because otherwise the process would not always converge to it regardless of the initial distribution.

8.4 Appendix D

In this appendix it will be presented the derivation of the kinetic energy term K of the local energy defined as

$$E_L(\mathbf{R}) = \frac{\langle \mathbf{R}|\hat{H}|\Psi\rangle}{\Psi(\mathbf{R})} \tag{8.46}$$

where $\mathbf{R}$ is a point in configuration space, and the wavefunction Ψ must be replaced with the appropriate function to approximate the real system. The goal is to compute explicitly the quantum force components that appear in K, that in coordinate representation is the sum of all second derivatives for all particles for both spin species:

$$K = -\frac{D}{\Psi}\left(\sum_i \sum_\alpha \partial^2_{\alpha_i}\Psi + \sum_a \sum_\alpha \partial^2_{\alpha_a}\Psi\right) \tag{8.47}$$

where the constant D is

$$D = \frac{\hbar^2}{2m} \tag{8.48}$$

and the mass of all particles are assumed equal.
If the system is dilute and made of bosonic particles, Ψ can be replaced with a Jastrow function, that is sufficient to describe correlations of few-body. The situation with dilute gas of fermions is different. Here a Jastrow function is not sufficient and an antisymmetrical component must be introduced. The Slater determinant is the simplest function to describe the nodal surface of the system.

Jastrow Function

A Jastrow function is the symmetrized product of a set of few-body functions that satisfy the cusp conditions. The easiest form contains only 2-body terms:

$$\Phi_J(\mathbf{R}) = \prod_{i<j} f(r_{ij})$$

where the product includes all couples of interacting particles (i, j), with the notation $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. If the short range approximation is valid, f can be the solution of 2-body interaction problem.
In case of two spin types, as for spin up and down $(\uparrow, \downarrow)$, we will have:

$$\Phi_J(\mathbf{R}) = J_{\uparrow\uparrow}(\mathbf{R})J_{\uparrow\downarrow}(\mathbf{R})J_{\downarrow\downarrow}(\mathbf{R}) = \prod_{i<j} f_{\uparrow\uparrow}(r_{ij}) \prod_{i,a} f_{\uparrow\downarrow}(r_{ia}) \prod_{a<b} f_{\downarrow\downarrow}(r_{ab}) \tag{8.49}$$

where we will always use in the following the label i for spin-up particles and a for spin-down particles. Let us define the Quantum Force for ϕ as:

$$F^\alpha_i = \frac{2}{\phi}\partial_{\alpha_i}\phi \tag{8.50}$$

where α is one of the three spatial directions (x, y, z). So the force components, for both spin species, are given by:

$$F_i^\alpha = \frac{2\partial_{\alpha_i}\Phi_J}{\Phi_J} = \frac{2\partial_{\alpha_i}(J_{\uparrow\uparrow}J_{\uparrow\downarrow}J_{\downarrow\downarrow})}{J_{\uparrow\uparrow}J_{\uparrow\downarrow}J_{\downarrow\downarrow}} = 2\left[\frac{\partial_{\alpha_i}J_{\uparrow\uparrow}}{J_{\uparrow\uparrow}} + \frac{\partial_{\alpha_i}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}}\right] = 2\left[\frac{\partial_{\alpha_i}\prod_{i<j} f_{\uparrow\uparrow}(r_{ij})}{\prod_{i<j} f_{\uparrow\uparrow}(r_{ij})} + \frac{\partial_{\alpha_i}\prod_{i,a} f_{\uparrow\downarrow}(r_{ia})}{\prod_{i,a} f_{\uparrow\downarrow}(r_{ia})}\right]$$
$$F_a^\alpha = \frac{2\partial_{\alpha_a}\Phi_J}{\Phi_J} = \frac{2\partial_{\alpha_a}(J_{\uparrow\uparrow}J_{\uparrow\downarrow}J_{\downarrow\downarrow})}{J_{\uparrow\uparrow}J_{\uparrow\downarrow}J_{\downarrow\downarrow}} = 2\left[\frac{\partial_{\alpha_a}J_{\uparrow\uparrow}}{J_{\uparrow\uparrow}} + \frac{\partial_{\alpha_a}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}}\right] = 2\left[\frac{\partial_{\alpha_a}\prod_{a<b} f_{\downarrow\downarrow}(r_{ab})}{\prod_{a<b} f_{\downarrow\downarrow}(r_{ab})} + \frac{\partial_{\alpha_a}\prod_{i,a} f_{\uparrow\downarrow}(r_{ia})}{\prod_{i,a} f_{\uparrow\downarrow}(r_{ia})}\right] \tag{8.51}$$

In these equations the partial derivative allows to rewrite the products of J functions as sums that simplify with the denominator up to the derivative of the function. Moreover the index of derivative i in first case, and a in second case, leads to equation where only this specific index for f is present, excluding components independent on the specific index.
Since the product of functions become a sum, where only one function per time is derivated, this will be the only element not simplified with the denominator. So for derivative with respect to spin up type: in the first term all functions without i index disappear, and a sum over different j appears. In the second case i is fixed while all possible values of a are present, and the sum is over these values. Similar is the procedure for spin down particles. So in conclusion the forces in both cases are:

$$\begin{aligned} F_i^\alpha &= 2\left[\sum_{j\neq i}\frac{\partial_{\alpha_i} f_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})} + \sum_a \frac{\partial_{\alpha_i} f_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right] \\ F_a^\alpha &= 2\left[\sum_{b\neq a}\frac{\partial_{\alpha_a} f_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})} + \sum_i \frac{\partial_{\alpha_a} f_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right] \end{aligned} \tag{8.52}$$

Now the function f must be derived with respect to direction α of i-th particle:

$$\begin{aligned} \partial_{\alpha_i} f_{\uparrow\uparrow}(r_{ij}) &= f'_{\uparrow\uparrow}(r_{ij})\partial_{\alpha_i}(|\mathbf{r}_i - \mathbf{r}_j|) = f'_{\uparrow\uparrow}(r_{ij})\frac{|\mathbf{r}_i - \mathbf{r}_j|}{(\mathbf{r}_i - \mathbf{r}_j)}\partial_{\alpha_i}(\mathbf{r}_i - \mathbf{r}_j) = \\ &= f'_{\uparrow\uparrow}(r_{ij})\frac{|\mathbf{r}_i - \mathbf{r}_j|}{|\mathbf{r}_i - \mathbf{r}_j|^2}(\mathbf{r}_i - \mathbf{r}_j)\cdot\hat{\alpha}_i = \frac{f'_{\uparrow\uparrow}(r_{ij})}{|\mathbf{r}_i - \mathbf{r}_j|}(\mathbf{r}_i - \mathbf{r}_j)\cdot\hat{\alpha}_i = f'_{\uparrow\uparrow}(r_{ij})\frac{(\alpha_i - \alpha_j)}{r_{ij}} \end{aligned} \tag{8.53}$$

where for the derivative $\mathbf{r}_i = \sum_\alpha \alpha_i\hat{\alpha}_i$ (and similarly also for $\mathbf{r}_j$ and so also for $\mathbf{r}_i - \mathbf{r}_j$). Then a vector 1 appears in position α_i and zero otherwise. In similar way one proceeds for mixed indices:

$$\begin{aligned} \partial_{\alpha_i} f_{\uparrow\downarrow}(r_{ia}) &= f'_{\uparrow\downarrow}(r_{ia})\partial_{\alpha_i}(|\mathbf{r}_i - \mathbf{r}_a|) = f'_{\uparrow\downarrow}(r_{ia})\frac{|\mathbf{r}_i - \mathbf{r}_a|}{(\mathbf{r}_i - \mathbf{r}_a)}\partial_{\alpha_i}(\mathbf{r}_i - \mathbf{r}_a) = \\ &= f'_{\uparrow\downarrow}(r_{ia})\frac{|\mathbf{r}_i - \mathbf{r}_a|}{|\mathbf{r}_i - \mathbf{r}_a|^2}(\mathbf{r}_i - \mathbf{r}_a)\cdot\hat{\alpha}_i = \frac{f'_{\uparrow\downarrow}(r_{ia})}{|\mathbf{r}_i - \mathbf{r}_a|}(\mathbf{r}_i - \mathbf{r}_a)\cdot\hat{\alpha}_i = f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i - \alpha_a)}{r_{ia}} \end{aligned} \tag{8.54}$$

Then the force is

$$F_i^\alpha = 2\left[\sum_{j\neq i}\frac{f'_{\uparrow\uparrow}(r_{ij})\frac{(\alpha_i-\alpha_j)}{r_{ij}}}{f_{\uparrow\uparrow}(r_{ij})} + \sum_a \frac{f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i-\alpha_a)}{r_{ia}}}{f_{\uparrow\downarrow}(r_{ia})}\right] \tag{8.55}$$

Analogue is the derivative of f with respect to spin down:

$$\partial_{\alpha_a} f_{\downarrow\downarrow}(r_{ab}) = f'_{\downarrow\downarrow}(r_{ab})\partial_{\alpha_a}(|\mathbf{r}_a - \mathbf{r}_b|) = f'_{\downarrow\downarrow}(r_{ab})\frac{|\mathbf{r}_a - \mathbf{r}_b|}{(\mathbf{r}_a - \mathbf{r}_b)}\partial_{\alpha_a}(\mathbf{r}_a - \mathbf{r}_b) =$$
$$= f'_{\downarrow\downarrow}(r_{ab})\frac{|\mathbf{r}_a - \mathbf{r}_b|}{|\mathbf{r}_a - \mathbf{r}_b|^2}(\mathbf{r}_a - \mathbf{r}_b)\cdot\hat{\alpha}_a = \frac{f'_{\downarrow\downarrow}(r_{ab})}{|\mathbf{r}_a - \mathbf{r}_b|}(\mathbf{r}_a - \mathbf{r}_b)\cdot\hat{\alpha}_a = f'_{\downarrow\downarrow}(r_{ab})\frac{(\alpha_a - \alpha_b)}{r_{ab}} \tag{8.56}$$

The only difference is a sign changed for the mixed indexes function

$$\partial_{\alpha_a} f_{\uparrow\downarrow}(r_{ia}) = f'_{\uparrow\downarrow}(r_{ia})\partial_{\alpha_a}(|\mathbf{r}_i - \mathbf{r}_a|) = f'_{\uparrow\downarrow}(r_{ia})\frac{|\mathbf{r}_i - \mathbf{r}_a|}{(\mathbf{r}_i - \mathbf{r}_a)}\partial_{\alpha_a}(\mathbf{r}_i - \mathbf{r}_a) =$$
$$= -f'_{\uparrow\downarrow}(r_{ia})\frac{|\mathbf{r}_i - \mathbf{r}_a|}{|\mathbf{r}_i - \mathbf{r}_a|^2}(\mathbf{r}_i - \mathbf{r}_a)\cdot\hat{\alpha}_a = -\frac{f'_{\uparrow\downarrow}(r_{ia})}{|\mathbf{r}_i - \mathbf{r}_a|}(\mathbf{r}_i - \mathbf{r}_a)\cdot\hat{\alpha}_a = -f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i - \alpha_a)}{r_{ia}} \tag{8.57}$$

So the force obtained in this case is:

$$F_a^{\alpha} = 2\left[\sum_{b\neq a}\frac{f_{\downarrow\downarrow}(r_{ab})\frac{(\alpha_a-\alpha_b)}{r_{ab}}}{f_{\downarrow\downarrow}(r_{ab})} - \sum_i \frac{f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i-\alpha_a)}{r_{ia}}}{f_{\uparrow\downarrow}(r_{ia})}\right] \tag{8.58}$$

This is the way to obtain the components for quantum force needed for K in eq.(8.47). So from definition let us substitute the wavefunction and develop all derivatives, then one can use previous results

$$K^J = -D\left[\sum_{i,\alpha}\left(\frac{\partial^2_{\alpha_i}J_{\uparrow\uparrow}}{J_{\uparrow\uparrow}} + \frac{\partial^2_{\alpha_i}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}} + \frac{\partial_{\alpha_i}J_{\uparrow\uparrow}\partial_{\alpha_i}J_{\uparrow\downarrow}}{J_{\uparrow\uparrow}J_{\uparrow\downarrow}}\right) + \sum_{a,\alpha}\left(\frac{\partial^2_{\alpha_a}J_{\downarrow\downarrow}}{J_{\downarrow\downarrow}} + \frac{\partial^2_{\alpha_a}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}} + \frac{\partial_{\alpha_a}J_{\downarrow\downarrow}\partial_{\alpha_a}J_{\uparrow\downarrow}}{J_{\downarrow\downarrow}J_{\uparrow\downarrow}}\right)\right] \tag{8.59}$$

as before, the derivatives of the products were developed and the terms that appeared both in numerator and denominator were simplified. Now, to calculate the derivatives of the function J, let us insert the definitions of the forces obtained above. One can start by evaluating the component referred to both spins up:

$$\partial^2_{\alpha_i}J_{\uparrow\uparrow} = \partial_{\alpha_i}\left(\partial_{\alpha_i}J_{\uparrow\uparrow}\right) = \partial_{\alpha_i}\left(\frac{J_{\uparrow\uparrow}F^{\alpha}_{i\uparrow\uparrow}}{2}\right)$$

let us replace the second derivative as:

$$K_{\uparrow\uparrow} = -D\sum_{i,\alpha}\frac{\partial^2_{\alpha_i}J_{\uparrow\uparrow}}{J_{\uparrow\uparrow}} = -D\sum_{i,\alpha}\frac{1}{J_{\uparrow\uparrow}}\partial_{\alpha_i}\left(\partial_{\alpha_i}J_{\uparrow\uparrow}\right) = -D\sum_{i,\alpha}\frac{1}{J_{\uparrow\uparrow}}\partial_{\alpha_i}\left(J_{\uparrow\uparrow}\sum_{j\neq i}\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\frac{(\alpha_i-\alpha_j)}{r_{ij}}\right) =$$
$$- D\sum_{i,\alpha}\left[\frac{\partial_{\alpha_i}J_{\uparrow\uparrow}}{J_{\uparrow\uparrow}}\sum_{j\neq i}\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\frac{(\alpha_i-\alpha_j)}{r_{ij}} + \partial_{\alpha_i}\sum_{j\neq i}\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\frac{(\alpha_i-\alpha_j)}{r_{ij}}\right] \tag{8.60}$$

recovering the definition of force in first term of last part, and carrying out the derivative in the second one:

$$K_{\uparrow\uparrow} = -D\sum_i\sum_\alpha^d\left\{\frac{(F^\alpha_{i\uparrow\uparrow})^2}{4}+\sum_{j\neq i}\left[\left(\frac{f''_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}-\left(\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\right)^2\right)\left(\frac{(\alpha_i-\alpha_j)}{r_{ij}}\right)^2\right]\right\}$$
$$-D\sum_i\sum_\alpha^d\left\{\sum_{j\neq i}\left[\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\left(\frac{1}{r_{ij}}-\frac{(\alpha_i-\alpha_j)^2}{r_{ij}^3}\right)\right]\right\}= \tag{8.61}$$
$$=-\frac{D}{4}\sum_i\sum_\alpha^d\left(F^\alpha_{i\uparrow\uparrow}\right)^2-D\sum_{i,\,j\neq i}\left[\frac{f''_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}-\left(\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\right)^2\right]-D\sum_{i,\,j\neq i}\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\frac{d-1}{r_{ij}}$$

where we used

$$\partial_{\alpha_i}\left(\frac{\alpha_i-\alpha_j}{|\mathbf{r}_i-\mathbf{r}_j|}\right)=\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}-\frac{\alpha_i-\alpha_j}{|\mathbf{r}_i-\mathbf{r}_j|^2}\frac{\mathbf{r}_i-\mathbf{r}_j}{|\mathbf{r}_i-\mathbf{r}_j|}\partial_{\alpha_i}(\mathbf{r}_i-\mathbf{r}_j)=\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}-\frac{(\alpha_i-\alpha_j)^2}{|\mathbf{r}_i-\mathbf{r}_j|^3} \tag{8.62}$$

The sum over α gives respectively $|\mathbf{r}_i-\mathbf{r}_j|$, $|\mathbf{r}_i-\mathbf{r}_j|^2$ and d (dimension of the problem) for α independent terms. This leads to the last equation.
A compact form can be written now, let us introduce a local kinetic energy term e_L (local kinetic energy of 2-body problem in units of $2D$) in $K_{\uparrow\uparrow}$, that becomes:

$$K_{\uparrow\uparrow}=-\frac{D}{4}\mathbf{F}_{\uparrow\uparrow}\cdot\mathbf{F}_{\uparrow\uparrow}+D\sum_{i,\,j\neq i}\left[e_L+\left(\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\right)^2\right] \tag{8.63}$$
$$e^L_{\uparrow\uparrow}=-\left(\frac{f''_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}+\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\frac{d-1}{r_{ij}}\right) \tag{8.64}$$

The other derivatives can be expressed similarly.
Let us turn now to both spin down components:

$$\partial^2_{\alpha_a}J_{\downarrow\downarrow}=\partial_{\alpha_a}\left(\partial_{\alpha_a}J_{\downarrow\downarrow}\right)=\partial_{\alpha_a}\left(\frac{J_{\downarrow\downarrow}F^\alpha_{a\downarrow\downarrow}}{2}\right)$$

so the second derivative is:

$$K_{\downarrow\downarrow}=-D\sum_{a,\alpha}\frac{\partial^2_{\alpha_a}J_{\downarrow\downarrow}}{J_{\downarrow\downarrow}}=-D\sum_{a,\alpha}\frac{1}{J_{\downarrow\downarrow}}\partial_{\alpha_a}\left(\partial_{\alpha_a}J_{\downarrow\downarrow}\right)=-D\sum_{a,\alpha}\frac{1}{J_{\downarrow\downarrow}}\partial_{\alpha_a}\left(J_{\downarrow\downarrow}\sum_{b\neq a}\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\frac{(\alpha_a-\alpha_b)}{r_{ab}}\right)=$$
$$-D\sum_{a,\alpha}\left[\frac{\partial_{\alpha_a}J_{\downarrow\downarrow}}{J_{\downarrow\downarrow}}\sum_{b\neq a}\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\frac{(\alpha_a-\alpha_b)}{r_{ab}}+\partial_{\alpha_a}\sum_{b\neq a}\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\frac{(\alpha_a-\alpha_b)}{r_{ab}}\right] \tag{8.65}$$

changing back the force definition in first term of last part, and performing the derivative of second term, we will have:

$$K_{\downarrow\downarrow}=-D\sum_a\sum_\alpha^d\left\{\frac{(F^\alpha_{a\downarrow\downarrow})^2}{4}+\sum_{b\neq a}\left[\left(\frac{f''_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}-\left(\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\right)^2\right)\left(\frac{(\alpha_a-\alpha_b)}{r_{ab}}\right)^2\right]\right\}$$
$$-D\sum_a\sum_\alpha^d\left\{\sum_{b\neq a}\left[\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\left(\frac{1}{r_{ab}}-\frac{(\alpha_a-\alpha_b)^2}{r_{ab}^3}\right)\right]\right\}= \tag{8.66}$$
$$=-\frac{D}{4}\sum_a\sum_\alpha^d\left(F^\alpha_{a\downarrow\downarrow}\right)^2-D\sum_{a,\,b\neq a}\left[\frac{f''_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}-\left(\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\right)^2\right]-D\sum_{a,\,b\neq a}\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\frac{d-1}{r_{ab}}$$

As before let us introduce the local kinetic energy e_L (local kinetic energy of 2-body porblem in unit of $2D$) to make more compact the equation for $K_{\downarrow\downarrow}$

$$K_{\downarrow\downarrow} = -\frac{D}{4}\mathbf{F}_{\downarrow\downarrow}\cdot\mathbf{F}_{\downarrow\downarrow} + D\sum_{a,\, b\neq a}\left[e^L_{\downarrow\downarrow} + \left(\frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\right)^2\right] \tag{8.67}$$

$$e^L_{\downarrow\downarrow} = -\left(\frac{f''_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})} + \frac{f'_{\downarrow\downarrow}(r_{ab})}{f_{\downarrow\downarrow}(r_{ab})}\frac{d-1}{r_{ab}}\right) \tag{8.68}$$

The terms dependent on both spin indexes are

$$\begin{aligned}\partial^2_{\alpha_i} J_{\uparrow\downarrow} &= \partial_{\alpha_i}\left(\partial_{\alpha_i} J_{\uparrow\downarrow}\right) = \partial_{\alpha_i}\left(\frac{J_{\uparrow\downarrow}F^{\alpha}_{i\uparrow\downarrow}}{2}\right)\\ \partial^2_{\alpha_a} J_{\uparrow\downarrow} &= \partial_{\alpha_a}\left(\partial_{\alpha_a} J_{\uparrow\downarrow}\right) = \partial_{\alpha_a}\left(\frac{J_{\uparrow\downarrow}F^{\alpha}_{a\uparrow\downarrow}}{2}\right)\end{aligned} \tag{8.69}$$

so, the second derivatives are:

$$\begin{aligned}K_{\uparrow\downarrow} =& -D\sum_{i,\alpha}\frac{\partial^2_{\alpha_i}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}} - D\sum_{a,\alpha}\frac{\partial^2_{\alpha_a}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}} = -D\sum_{i,\alpha}\frac{\partial_{\alpha_i}\left(J_{\uparrow\downarrow}F^{\alpha}_{i\uparrow\downarrow}\right)}{2J_{\uparrow\downarrow}} - D\sum_{a,\alpha}\frac{\partial_{\alpha_a}\left(J_{\uparrow\downarrow}F^{\alpha}_{a\uparrow\downarrow}\right)}{2J_{\uparrow\downarrow}} =\\ =& -D\sum_{i,\alpha}\frac{1}{J_{\uparrow\downarrow}}\partial_{\alpha_i}\left(J_{\uparrow\downarrow}\sum_a f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|}\right) + D\sum_{a,\alpha}\frac{1}{J_{\uparrow\downarrow}}\partial_{\alpha_a}\left(J_{\uparrow\downarrow}\sum_i f'_{\uparrow\downarrow}(r_{ia})\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|}\right) =\\ =& -D\sum_{i,\alpha}\left[\frac{\partial_{\alpha_i}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}}\sum_a\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|} + \partial_{\alpha_i}\sum_a\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|}\right]\\ &+ D\sum_{a,\alpha}\left[\frac{\partial_{\alpha_a}J_{\uparrow\downarrow}}{J_{\uparrow\downarrow}}\sum_i\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|} + \partial_{\alpha_a}\sum_i\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{(\alpha_i-\alpha_a)}{|r_i-r_a|}\right]\end{aligned} \tag{8.70}$$

changing back the force definition, and performing the derivative of second term, we will have:

$$\begin{aligned}K_{\uparrow\downarrow} =& -D\left(\sum_i\sum_\alpha^d\frac{(F^{\alpha}_{i\uparrow\downarrow})^2}{4} + \sum_a\sum_\alpha^d\frac{(F^{\alpha}_{a\uparrow\downarrow})^2}{4}\right)\\ &- D\sum_i\sum_\alpha^d\sum_a\left[\left(\frac{f''_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})} - \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2\right)\left(\frac{(\alpha_i-\alpha_a)}{r_{ia}}\right)^2 + \frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\left(\frac{1}{r_{ia}} - \frac{(\alpha_i-\alpha_a)^2}{r^3_{ia}}\right)\right]\\ &- D\sum_a\sum_\alpha^d\sum_i\left[\left(\frac{f''_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})} - \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2\right)\left(\frac{(\alpha_i-\alpha_a)}{r_{ia}}\right)^2 + \frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\left(\frac{1}{r_{ia}} - \frac{(\alpha_i-\alpha_a)^2}{r^3_{ia}}\right)\right] =\\ =& -\frac{D}{4}\mathbf{F}_A\cdot\mathbf{F}_A - 2D\sum_{i\,a}\left[\frac{f''_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})} - \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2 + \frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{d-1}{r_{ia}}\right]\end{aligned} \tag{8.71}$$

As before let us introduce the local kinetic energy e_L (local kinetic energy of 2-body porblem in units of $2D$) to make more compact the equation for $K_{\downarrow\downarrow}$

$$K_{\uparrow\downarrow} = -\frac{D}{4}\mathbf{F}_A\cdot\mathbf{F}_A + 2D\sum_{i\,a}\left[e^L_{\uparrow\downarrow} + \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2\right] \tag{8.72}$$

$$e^L_{\uparrow\downarrow} = -\left(\frac{f''_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})} + \frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\frac{d-1}{r_{ia}}\right) \tag{8.73}$$

As last thing let us compute the first derivative term

$$\begin{aligned}
&- D\sum_i\sum_\alpha^d 2\frac{\partial_{\alpha_i} J_{\uparrow\uparrow}\partial_{\alpha_i} J_{\uparrow\downarrow}}{J_{\uparrow\uparrow}J_{\uparrow\downarrow}} - D\sum_a\sum_\alpha^d 2\frac{\partial_{\alpha_a} J_{\downarrow\downarrow}\partial_{\alpha_a} J_{\uparrow\downarrow}}{J_{\downarrow\downarrow}J_{\uparrow\downarrow}} \\
&= -D\sum_i\sum_\alpha^d 2\frac{F^\alpha_{i\uparrow\uparrow}}{2}\cdot\frac{F^\alpha_{i\uparrow\downarrow}}{2} - D\sum_a\sum_\alpha^d 2\frac{F^\alpha_{a\uparrow\uparrow}}{2}\cdot\frac{F^\alpha_{a\uparrow\downarrow}}{2} \\
&= -\frac{D}{2}\left(F_{\uparrow\uparrow}\cdot F_{\uparrow\downarrow} + F_{\uparrow\uparrow}\cdot F_{\uparrow\downarrow}\right) = -\frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A
\end{aligned} \tag{8.74}$$

Finally all expressions together in K^J lead to

$$\begin{aligned}
K^J = &-\frac{D}{4}\mathbf{F}_{\uparrow\uparrow}\cdot\mathbf{F}_{\uparrow\uparrow} + D\sum_{i,\,j\neq i}\left[e^L_{\uparrow\uparrow}(r_{ij}) + \left(\frac{f'_{\uparrow\uparrow}(r_{ij})}{f_{\uparrow\uparrow}(r_{ij})}\right)^2\right] - \frac{D}{4}\mathbf{F}_{\downarrow\downarrow}\cdot\mathbf{F}_{\downarrow\downarrow} + D\sum_{a,\,b\neq a}\left[e^L_{\downarrow\downarrow}(r_{ab}) + \left(\frac{f'_{\uparrow\uparrow}(r_{ab})}{f_{\uparrow\uparrow}(r_{ab})}\right)^2\right] \\
&- \frac{D}{4}\mathbf{F}_A\cdot\mathbf{F}_A + 2D\sum_{i,a}\left[e^L_{\uparrow\downarrow}(r_{ia}) + \left(\frac{f'_{\uparrow\downarrow}(r_{ia})}{f_{\uparrow\downarrow}(r_{ia})}\right)^2\right] - \frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A = K_{\uparrow\uparrow} + K_{\uparrow\downarrow} + K_{\downarrow\downarrow} - \frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A
\end{aligned} \tag{8.75}$$

where $\mathbf{F}_S = \left\{F^\alpha_{i\uparrow\uparrow},\ F^\alpha_{a\downarrow\downarrow}\right\}$, $\mathbf{F}_A = \left\{F^\alpha_{i\uparrow\downarrow},\ F^\alpha_{a\uparrow\downarrow}\right\}$.

Slater Determinant

Let us move now to the fermionic case, where the sole Jastrow function is not enough to describe correctly the system, and the antisymmetry must be considered. The easiest function is the one constructed adiabatically from non-interacting wavefunction.
For a correct DMC simulation a good nodal surface is fundamental, while a wavefunction optimization is not necessary, however it has effect over the variance. In case of diluted system a Slater determinant is sufficient. It is constructed from orbitals of single non-interacting particles $\phi_{\mathbf{k}_p}(\mathbf{r}_i)$. So plane waves can describe the system.

Since the system under examination is composed of both spin up and down, the function will be composed of one Jastrow function and two Slater determinants (D): $\Psi^{JS} = \Phi_J D_\uparrow D_\downarrow$, where $D_\uparrow(\mathbf{R}) = det\, S_{\uparrow\, pi} = det\left[\phi_{\mathbf{k}_p}(\mathbf{r}_i)\right]$. Inside the determinant D, single orbitals are expressed as combination of exponentials, then exponentials can be replaced by sin and cos functions. This is a more convenient form when numerical calculation are performed, because they allow for using a purely real wavefunction.
In order to describe determinant's properties, we can simplify the notation and drop the spin indexes, the result will be valid for both species. Let us introduce the inverse S matrix:

$$\sum_i S_{pi}\bar{S}_{iq} = \delta_{pq},\ \sum_p \bar{S}_{jp}S_{pi} = \delta_{ij}$$

It is possible to write the determinant by means of cofactor matrix as $D = \sum_i S_{pi}A_{ip}$, and also $\bar{S}_{ip} = A_{ip}/D$. These definitions lead to some properties useful to compute the derivatives of the determinant that appear in kinetic energy of Ψ^{JS}

$$\begin{aligned}
\frac{1}{D}\frac{\partial D}{\partial S_{pi}} &= \bar{S}_{ip} \\
\frac{\partial \bar{S}_{ip}}{\partial S_{qj}} &= -\bar{S}_{iq}\bar{S}_{jp}
\end{aligned} \tag{8.76}$$

We can prove the first equation starting from the definition of the determinant. It is a polynomial function and, therefore, it is expressible as a function of the matrix elements, and then differentiable with chain rule: $D = F(S_{11}, \cdots, S_{nn})$, $\det : \mathbf{R}^{n\times n} \to \mathbf{R}$, so the differential will be $dD = \sum_{pi} \frac{\partial D}{\partial S_{pi}} dS_{pi}$. In the definition of the determinant with the cofactor matrix the sum index is arbitrary, both as a row or column index, so we can use the same index i as the in the derivative rewriting

$$\begin{aligned}\frac{\partial D}{\partial S_{qi}} &= \frac{\partial \sum_p S_{pi} A_{ip}}{\partial S_{qi}} = \sum_p \frac{\partial}{\partial S_{qi}}(S_{pi} A_{ip}) \\ &= \sum_p \left(\frac{\partial S_{pi}}{\partial S_{qi}} A_{ip} + S_{pi} \frac{\partial A_{ip}}{\partial S_{qi}} \right) = \sum_p A_{ip} \frac{\partial S_{pi}}{\partial S_{qi}} = \sum_p A_{ip} \delta_{pq} = A_{iq}\end{aligned} \tag{8.77}$$

Where we used the fact that the cofactor matrix has elements independent of the row and column of the matrix from which it was calculated, therefore the derivative with respect to a matrix element with the same row or column index will be zero. Moreover when we calculate the derivative of a matrix element with respect to another, we obtain a Kronecker delta, as they are independent of each other. Then $\frac{\partial D}{\partial S_{pi}} = A_{ip}$, and, exploiting the definition of inverse matrix $\bar{S}_{ip} = A_{ip}/D$ and dividing both by D, the first property is proved.
To obtain the second of the two properties, we can start by considering the following derivative

$$\frac{\partial}{\partial S_{qj}} S \times \bar{S} = \frac{\partial}{\partial S_{qj}} \mathcal{I} = 0 \tag{8.78}$$

that can be developed as

$$\frac{\partial}{\partial S_{qj}} S \times \bar{S} = \frac{\partial S}{\partial S_{qj}} \times \bar{S} + S \times \frac{\partial \bar{S}}{\partial S_{qj}} = 0 \tag{8.79}$$

moving to the right side the first left term, and multiplying both sides by the inverse matrix we have

$$\frac{\partial \bar{S}}{\partial S_{qj}} = -\bar{S} \times \frac{\partial S}{\partial S_{qj}} \times \bar{S} \tag{8.80}$$

now if we write explicitly the spin indexes

$$\frac{\partial \bar{S}_{ip}}{\partial S_{qj}} = -\bar{S}_{ir} \delta_{rq} \delta_{il} \bar{S}_{lp} = -\bar{S}_{iq} \bar{S}_{ip} \tag{8.81}$$

With these two properties we can proceed in the calculation of the quantum forces for the Slater determinant necessary to obtain the kinetic energy K^S and that of the compound function K^{JS}. Let us first consider the case of spin up, knowing that to obtain the analogous quantities for spin down it will be enough to substitute the derivative indices,

$$F_i^\alpha = \frac{2}{D_\uparrow} \partial_{\alpha_i} D_\uparrow = \frac{2}{D_\uparrow} \sum_p \frac{\partial D_\uparrow}{S_{ip}} \frac{\partial S_{ip}}{\partial \alpha_i} = 2 \sum_p \bar{S}_{pi} \partial_{\alpha_i} S_{ip} \tag{8.82}$$

this force will be used for the calculation of the second derivative in K

$$\begin{aligned}K^S &= -\frac{D}{D_\uparrow} \sum_i \sum_\alpha^d \partial_{\alpha_i}^2 D_\uparrow = -\frac{D}{D_\uparrow} \sum_i \sum_\alpha^d \partial_{\alpha_i} (\partial_{\alpha_i} D_\uparrow) = -\frac{D}{D_\uparrow} \sum_i \sum_\alpha^d \partial_{\alpha_i} (D_\uparrow \sum_p \bar{S}_{pi} \partial_{\alpha_i} S_{ip}) \\ &= -D \sum_i \sum_\alpha^d \left[\frac{\partial_{\alpha_i} D_\uparrow}{D_\uparrow} \sum_p \bar{S}_{pi} \partial_{\alpha_i} S_{ip} + \sum_p \left(\frac{\partial \bar{S}_{pi}}{\partial \alpha_i} \partial_{\alpha_i} S_{ip} + \bar{S}_{pi} \partial_{\alpha_i}^2 S_{ip} \right) \right]\end{aligned} \tag{8.83}$$

In the first term of the last equality we can replace the definition of force, while in the second term the second of the properties in eq(8.76), so we obtain

$$K_{\uparrow}^{S} = -D\sum_{i}\sum_{\alpha}^{d}\left[\sum_{pq}\bar{S}_{qi}\partial_{\alpha_i}S_{iq}\bar{S}_{pi}\partial_{\alpha_i}S_{ip} - \sum_{pq}\bar{S}_{qi}\bar{S}_{pi}\partial_{\alpha_i}S_{iq}\partial_{\alpha_i}S_{ip} + \sum_{p}\bar{S}_{pi}\partial_{\alpha_i}^{2}S_{ip}\right]$$
$$= -D\sum_{i}\sum_{\alpha}^{d}\sum_{p}\bar{S}_{pi}\partial_{\alpha_i}^{2}S_{ip} \quad (8.84)$$

In an analogue way we can obtain for spin down

$$K_{\downarrow}^{S} = -D\sum_{a}\sum_{\alpha}^{d}\left[\sum_{pq}\bar{S}_{qa}\partial_{\alpha_a}S_{aq}\bar{S}_{pa}\partial_{\alpha_a}S_{ap} - \sum_{pq}\bar{S}_{qa}\bar{S}_{pa}\partial_{\alpha_a}S_{aq}\partial_{\alpha_a}S_{ap} + \sum_{p}\bar{S}_{pa}\partial_{\alpha_a}^{2}S_{ap}\right]$$
$$= -D\sum_{a}\sum_{\alpha}^{d}\sum_{p}\bar{S}_{pa}\partial_{\alpha_a}^{2}S_{ap} \quad (8.85)$$

When the matrix elements are plane waves

$$S_{pi} = e^{i\mathbf{k}_p\cdot\mathbf{r}_i} \quad (8.86)$$

we have the following simplification in all kinetic energy components

$$\bar{S}_{ip}\partial_{\alpha_i}^{2}S_{pi} = \bar{S}_{ip}\left[-(\mathbf{k}_p^{\alpha})^2 S_{pi}\right] = -(\mathbf{k}_p^{\alpha})^2 \quad (8.87)$$

So the kinetic energy reduces to sum of squared wavevectors for different values of p.
Let us compute now K^{JS} for the total wavefunction $\Phi^{JS} = \Phi_J D_{\uparrow} D_{\downarrow}$, this function leads to

$$K^{JS} = -\frac{D}{\Phi^{JS}}\left(\sum_{i}\sum_{\alpha}^{d}\partial_{\alpha_i}^{2}\Phi^{JS} + \sum_{a}\sum_{\alpha}^{d}\partial_{\alpha_a}^{2}\Phi^{JS}\right) \quad (8.88)$$

For the first term one has

$$-\frac{D}{\Phi^{JS}}\sum_{i}\sum_{\alpha}^{d}\partial_{\alpha_i}^{2}\Phi^{JS} = -\frac{D}{\Phi_J D_{\uparrow} D_{\downarrow}}\sum_{i}\sum_{\alpha}^{d}\partial_{\alpha_i}^{2}\left(\Phi_J D_{\uparrow} D_{\downarrow}\right)$$
$$= -\frac{D}{\Phi_J D_{\uparrow} D_{\downarrow}}\sum_{i}\sum_{\alpha}^{d}\partial_{\alpha_i}\left(\partial_{\alpha_i}\Phi_J D_{\uparrow} D_{\downarrow} + \Phi_J\partial_{\alpha_i}D_{\uparrow} D_{\downarrow}\right) = \quad (8.89)$$
$$-D\sum_{i}\sum_{\alpha}^{d}\left(\frac{\partial_{\alpha_i}^{2}\Phi_J}{\Phi_J} + \frac{\partial_{\alpha_i}^{2}D_{\uparrow}}{D_{\uparrow}} + 2\frac{\partial_{\alpha_i}\Phi_J\partial_{\alpha_i}D_{\uparrow}}{\Phi_J D_{\uparrow}}\right)$$

since $D_{\downarrow}$ is independent on the spin up index i. Analogue is the derivative with respect to the index a

$$-\frac{D}{\Phi^{JS}}\sum_{a}\sum_{\alpha}^{d}\partial_{\alpha_i}^{2}\Phi^{JS} = -D\sum_{a}\sum_{\alpha}^{d}\left(\frac{\partial_{\alpha_a}^{2}\Phi_J}{\Phi_J} + \frac{\partial_{\alpha_a}^{2}D_{\downarrow}}{D_{\downarrow}} + 2\frac{\partial_{\alpha_a}\Phi_J\partial_{\alpha_a}D_{\downarrow}}{\Phi_J D_{\downarrow}}\right) \quad (8.90)$$

Let us substitute the above results and that of previous section

$$
\begin{aligned}
K^J &= -\frac{D}{\Phi_J}\left(\sum_i\sum_\alpha \partial^2_{\alpha_i}\Phi_J + \sum_a\sum_\alpha \partial^2_{\alpha_a}\Phi_J\right) = K_{\uparrow\uparrow} + K_{\uparrow\downarrow} + K_{\downarrow\downarrow} - \frac{D}{2}\mathbf{F}_S\cdot\mathbf{F}_A \\
K^S_\uparrow &= -\frac{D}{D_\uparrow}\sum_i\sum_\alpha^d \partial^2_{\alpha_i}D_\uparrow = -D\sum_i\sum_\alpha^d\sum_p \bar{S}_{pi}\partial^2_{\alpha_i}S_{ip} \\
K^S_\downarrow &= -\frac{D}{D_\downarrow}\sum_a\sum_\alpha^d \partial^2_{\alpha_a}D_\downarrow = -D\sum_a\sum_\alpha^d\sum_p \bar{S}_{pa}\partial^2_{\alpha_a}S_{ap}
\end{aligned}
\tag{8.91}
$$

Finally let us put everything together in K^{JS}

$$
\begin{aligned}
K^{JS} &= -D\sum_i\sum_\alpha^d\left(\frac{\partial^2_{\alpha_i}\Phi_J}{\Phi_J} + \frac{\partial^2_{\alpha_i}D_\uparrow}{D_\uparrow} + 2\frac{\partial_{\alpha_i}\Phi_J\partial_{\alpha_i}D_\uparrow}{\Phi_J D_\uparrow}\right) \\
&\quad - D\sum_a\sum_\alpha^d\left(\frac{\partial^2_{\alpha_a}\Phi_J}{\Phi_J} + \frac{\partial^2_{\alpha_a}D_\downarrow}{D_\downarrow} + 2\frac{\partial_{\alpha_a}\Phi_J\partial_{\alpha_a}D_\downarrow}{\Phi_J D_\downarrow}\right) \\
&= -D\left(\sum_i\sum_\alpha^d\frac{\partial^2_{\alpha_i}\Phi_J}{\Phi_J} + \sum_a\sum_\alpha^d\frac{\partial^2_{\alpha_a}\Phi_J}{\Phi_J}\right) - D\sum_i\sum_\alpha^d\frac{\partial^2_{\alpha_i}D_\uparrow}{D_\uparrow} - D\sum_a\sum_\alpha^d\frac{\partial^2_{\alpha_a}D_\downarrow}{D_\downarrow} \\
&\quad - D\left(\sum_i\sum_\alpha^d 2\frac{\partial_{\alpha_i}\Phi_J\partial_{\alpha_i}D_\uparrow}{\Phi_J D_\uparrow} + \sum_a\sum_\alpha^d 2\frac{\partial_{\alpha_a}\Phi_J\partial_{\alpha_a}D_\downarrow}{\Phi_J D_\downarrow}\right) \\
&= K^J + K^S_\uparrow + K^S_\downarrow - 2D\left[\sum_{i\,\alpha}\frac{\partial_{\alpha_i}\Phi_J}{\Phi_J}\frac{\partial_{\alpha_i}D_\uparrow}{D_\uparrow} + \sum_{a\,\alpha}\frac{\partial_{\alpha_a}\Phi_J}{\Phi_J}\frac{\partial_{\alpha_a}D_\downarrow}{D_\downarrow}\right]
\end{aligned}
\tag{8.92}
$$

Let us replace now the definition of quantum force inside the previous equation to obtain a more compact form

$$
\begin{aligned}
K^{JS} &= K^J + K^S_\uparrow + K^S_\downarrow - \frac{D}{2}\left[\sum_{i\,\alpha}F_i^{J\,\alpha}\cdot F_i^{S\,\alpha} + \sum_{a\,\alpha}F_a^{J\,\alpha}\cdot F_a^{S\,\alpha}\right] \\
&= K^J + K^S_\uparrow + K^S_\downarrow - \frac{D}{2}\mathbf{F}^J\cdot\mathbf{F}^S
\end{aligned}
\tag{8.93}
$$

BIBLIOGRAPHY

[1] M. Pini, P. Pieri, and G. Calvanese Strinati (Unpublished).

[2] P. Fulde and R. A. Ferrell, Physical Review **135**, A550 (1964).

[3] A. Larkin and I. Ovchinnikov, Soviet Physics-JETP **20**, 762 (1965).

[4] Y. Matsuda and H. Shimahara, Journal of the Physical Society of Japan **76**, 051005 (2007).

[5] R. Lortz, Y. Wang, A. Demuer, P. H. M. Böttger, B. Bergk, G. Zwicknagl, Y. Nakazawa, and J. Wosnitza, Physical Review Letters **99**, 187002 (2007).

[6] C. C. Agosta, J. Jin, W. A. Coniglio, B. E. Smith, K. Cho, I. Stroe, C. Martin, S. W. Tozer, T. P. Murphy, E. C. Palm, et al., Physical Review B **85**, 214514 (2012).

[7] G. Koutroulakis, H. Kühne, J. A. Schlueter, J. Wosnitza, and S. E. Brown, Physical Review Letters **116**, 067003 (2016).

[8] C.-w. Cho, J. H. Yang, N. F. Q. Yuan, J. Shen, T. Wolf, and R. Lortz, Physical Review Letters **119**, 217002 (2017).

[9] J. Wosnitza, Annalen der Physik **530**, 1700282 (2018).

[10] K. Kinjo, M. Manago, S. Kitagawa, Z. Mao, S. Yonezawa, Y. Maeno, and K. Ishida, Science **376**, 397 (2022).

[11] L. D. Landau, E. Lifshitz, and L. Pitaevskii, *Course of Theoretical Physics: Statistical Physics, Part 2: by EM Lifshitz and LP Pitaevskii* (Elsevier, 2013).

[12] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of quantum field theory in statistical physics* (Courier Corporation, 2012).

[13] L. D. Landau, Soviet Physics Jetp-Ussr **3**, 920 (1957).

[14] L. Landau, Soviet Physics Jetp-Ussr **5**, 101 (1957).

[15] L. Landau, Sov. Phys. JETP **8**, 70 (1959).

[16] J. Luttinger, Physical Review **119**, 1153 (1960).

[17] P. Coleman, *Introduction to Many-Body Physics* (Cambridge University Press, 2015).

[18] I. I. Pomeranchuk et al., Sov. Phys. JETP **8**, 361 (1958).

[19] N. F. Mott, Proceedings of the Physical Society. Section A **62**, 416 (1949).

[20] G. D. Mahan, *Many-particle physics* (Kluwer Academic/Plenum Publishers, 2000).

[21] P. Nozieres, *Theory of interacting Fermi systems* (CRC Press, 2018).

[22] K. Quader and K. Bedell, Journal of low temperature physics **58**, 89 (1985).

[23] A. Schwenk, G. E. Brown, and B. Friman, Nuclear Physics A **703**, 745 (2002).

[24] O. Sjöberg, Nuclear Physics A **265**, 511 (1976).

[25] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Physical Review **108**, 1175 (1957).

[26] K. Maki, Physical Review **148**, 362 (1966).

[27] J. J. Kinnunen, J. E. Baarsma, J.-P. Martikainen, and P. Törmä, Reports on Progress in Physics **81**, 046401 (2018).

[28] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Science **269**, 198 (1995).

[29] C. C. Bradley, C. Sackett, J. Tollett, and R. G. Hulet, Physical Review Letters **75**, 1687 (1995).

[30] B. DeMarco, J. Bohn, J. Burke Jr, M. Holland, and D. S. Jin, Physical Review Letters **82**, 4208 (1999).

[31] F. Schreck, L. Khaykovich, K. Corwin, G. Ferrari, T. Bourdel, J. Cubizolles, and C. Salomon, Physical Review Letters **87**, 080403 (2001).

[32] S. Giorgini, L. P. Pitaevskii, and S. Stringari, Reviews of Modern Physics **80**, 1215 (2008).

[33] H. T. Stoof, K. B. Gubbels, and D. B. Dickerscheid, *Ultracold Quantum Fields* (Springer, 2009).

[34] F. Schreck and K. v. Druten, Nature Physics **17**, 1296 (2021).

[35] A. L. Migdall, J. V. Prodan, W. D. Phillips, T. H. Bergeman, and H. J. Metcalf, Physical Review Letters **54**, 2596 (1985).

[36] H. F. Hess, Physical Review B **34**, 3476 (1986).

[37] J. Billy, V. Josse, Z. Zuo, A. Bernard, B. Hambrecht, P. Lugan, D. Clément, L. Sanchez-Palencia, P. Bouyer, and A. Aspect, Nature **453**, 891 (2008).

[38] H. Feshbach, Annals of Physics **5**, 357 (1958).

[39] G. C. Strinati, P. Pieri, G. Röpke, P. Schuck, and M. Urban, Physics Reports **738**, 1 (2018).

[40] W. Zwerger, *The BCS-BEC crossover and the unitary Fermi gas*, vol. 836 (Springer Science & Business Media, 2011).

[41] P. S. Jessen, C. Gerz, P. D. Lett, W. D. Phillips, S. Rolston, R. Spreeuw, and C. Westbrook, Physical Review Letters **69**, 49 (1992).

[42] M. W. Zwierlein, A. Schirotzek, C. H. Schunck, and W. Ketterle, Science **311**, 492 (2006).

[43] J. Stenger, S. Inouye, M. Andrews, H.-J. Miesner, D. Stamper-Kurn, and W. Ketterle, Physical Review Letters **82**, 2422 (1999).

[44] D. Petrov, C. Salomon, and G. V. Shlyapnikov, Physical Review Letters **93**, 090404 (2004).

[45] Z. Roberto, *Quantum Mechanics: Lecture Notes* (Dipartimento di Fisica ed Astronomia, Università di Bologna, 2022).

[46] S. Inouye, M. Andrews, J. Stenger, H.-J. Miesner, D. M. Stamper-Kurn, and W. Ketterle, Nature **392**, 151 (1998).

[47] K. B. Gubbels and H. T. Stoof, Physics Reports **525**, 255 (2013).

[48] S. Simonucci, P. Pieri, and G. Strinati, EPL (Europhysics Letters) **69**, 713 (2005).

[49] M. Romans and H. Stoof, Physical Review A **74**, 053618 (2006).

[50] C. Lobo, A. Recati, S. Giorgini, and S. Stringari, Physical Review Letters **97**, 200403 (2006).

[51] S. Nascimbène, N. Navon, S. Pilati, F. Chevy, S. Giorgini, A. Georges, and C. Salomon, Physical Review Letters **106**, 215303 (2011).

[52] L. Landau, Phys. Z. Sowjetunion **3**, 644 (1933).

[53] S. Pekar, Zh. Eksp. Teor. Fiz **16**, 335 (1946).

[54] H. Fröhlich, H. Pelzer, and S. Zienau, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **41**, 221 (1950).

[55] P. Massignan, M. Zaccanti, and G. M. Bruun, Reports on Progress in Physics **77**, 034401 (2014).

[56] F. Chevy, Physical Review A **74**, 063628 (2006).

[57] L. Radzihovsky and D. E. Sheehy, Reports on Progress in Physics **73**, 076501 (2010).

[58] Y.-i. Shin, C. H. Schunck, A. Schirotzek, and W. Ketterle, Nature **451**, 689 (2008).

[59] Y.-i. Shin, A. Schirotzek, C. H. Schunck, and W. Ketterle, Physics Review Letters **101**, 070404 (2008).

[60] B. A. Olsen, M. C. Revelle, J. A. Fry, D. E. Sheehy, and R. G. Hulet, Physics Review A **92**, 063616 (2015).

[61] S. Pilati and S. Giorgini, Physical Review Letters **100**, 030401 (2008).

[62] A. Bulgac and M. M. Forbes, Physics Review Letters **101**, 215301 (2008).

[63] W. L. McMillan, Physical Review **138**, A442 (1965).

[64] M. E. Newman and G. T. Barkema, *Monte Carlo methods in statistical physics* (Clarendon Press, 1999).

[65] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods* (John Wiley & Sons, 2009).

[66] N. Metropolis and S. Ulam, Journal of the American statistical association **44**, 335 (1949).

[67] H. Dörrie, *100 great problems of elementary mathematics* (Courier Corporation, 2013).

[68] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, The journal of chemical physics **21**, 1087 (1953).

[69] D. Ceperley, G. V. Chester, and M. H. Kalos, Physical Review B **16**, 3081 (1977).

[70] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, 2016).

[71] F. James, *Statistical Methods in Experimental Physics* (World Scientific, 2006), 2nd ed.

[72] C. Fourgeaud and A. Fuchs, Zbl0201 **51501** (1967).

[73] G. E. Box, Annals of Mathematical Statistics **29**, 610 (1958).

[74] C. J. Thompson, *Mathematical statistical mechanics* (Princeton University Press, 2015).

[75] C. D. Meyer, *Matrix analysis and applied linear algebra*, vol. 71 (Siam, 2000).

[76] C. R. MacCluer, SIAM Review **42**, 487 (2000).

[77] W. Wood and F. Parker, The Journal of Chemical Physics **27**, 720 (1957).

[78] M. Rosenbluth, Tech. Rep., Los Alamos Scientific Lab. (1953).

[79] W. K. Hastings, Biometrika **57**, 97 (1970).

[80] W. Feller, *An introduction to probability theory and its applications, vol 2* (John Wiley & Sons, 2008).

[81] G. Bertaina, Ph.D. thesis, University of Trento (2010).

[82] J. Thijssen, *Computational physics* (Cambridge university press, 2007).

[83] B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo methods in ab initio quantum chemistry*, vol. 1 (World Scientific, 1994).

[84] R. Guardiola, in *Microscopic quantum many-body theories and their applications* (Springer, 1998), pp. 269–336.

[85] M. H. Kalos, D. Levesque, and L. Verlet, Physical Review A **9**, 2178 (1974).

[86] P. J. Reynolds, J. Tobochnik, and H. Gould, Computers in Physics **4**, 662 (1990).

[87] M. Troyer and U.-J. Wiese, Physical Review Letters **94**, 170201 (2005).

[88] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester Jr, The Journal of Chemical Physics **77**, 5593 (1982).

[89] J. Toulouse, R. Assaraf, and C. J. Umrigar, in *Advances in Quantum Chemistry* (Elsevier, 2016), vol. 73, pp. 285–314.

[90] S. Giorgini, J. Boronat, and J. Casulleras, Physical Review A **60**, 5129 (1999).

[91] D. Ceperley, G. V. Chester, and M. H. Kalos, Physical Review B **16**, 3081 (1977).

[92] J. Bouchaud, A. Georges, and C. Lhuillier, Journal de Physique **49**, 553 (1988).

[93] G. Astrakharchik, J. Boronat, J. Casulleras, and S. Giorgini, Physical Review Letters **93**, 200404 (2004).

[94] F. Arias de Saavedra, J. Boronat, A. Polls, and A. Fabrocini, Physical Review B **50**, 4248 (1994).

[95] M. F. DePasquale, S. M. Rothstein, and J. Vrbik, The Journal of Chemical Pjysics **89**, 3629 (1988).

[96] S. Nascimbene, N. Navon, K. Jiang, L. Tarruell, M. Teichmann, J. Mckeever, F. Chevy, and C. Salomon, Physical Review Letters **103**, 170402 (2009).

[97] C. Lin, F. Zong, and D. M. Ceperley, Physical Review E **64**, 016702 (2001).