Towards Improving Configuration Path Integral Monte Carlo

Bachelorarbeit
im Bachelorstudiengang Physik (1-Fach)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel
vorgelegt von

Paul Luis Röhl

Erstgutachter: Prof. Dr. Michael Bonitz
Zweitgutachter: Dr. Alexey Filinov

Kiel im September 2021

Zusammenfassung

In der Plasmaphysik sind Modelle warmer dichter Materie, wie das homogene Elektronengas, von zunehmender Wichtigkeit. Durch Fehler bei analytischen Näherungen und Schwierigkeiten bei theoretischen Modellen, zum Beispiel die Berücksichtigung der Coulomb-Wechselwirkung aller Elektronen im homogenen Elektronengas oder das aus der Spinstatistik resultierende Pauli-Prinzip, gibt es einen Bedarf an *ab initio* Simulationen der thermodynamischen Eigenschaften des Elektronengases unter verschiedenen Temperaturen und Kopplungsstärken.

Eine Möglichkeit, solche durchzuführen, sind auf Feynmans Pfadintegralformalismus basierende Quanten-Monte Carlo-Techniken. Diese sind allerdings stark von dem fermionischen Vorzeichenproblem beeinträchtigt, welches teilweise durch die Verwendung verschiedener komplementärer Formulierungen umgangen werden kann. Eine dieser Formulierungen ist durch den in zweiter Quantisierung beschriebenen Configuration Path Integral Monte Carlo Algorithmus gegeben.

In dieser Bachelorarbeit werden zwei Ideen ausgearbeitet, deren Ziel die Verbesserung dieses Algorithmus ist. Der erste Teil untersucht Propagatoren höherer Ordnung bei der Herleitung der thermodynamischen Zustandssumme und zeigt, wie man kanonisch diskrete Simulationen beliebiger Ordnung durchführen kann. Im zweiten Teil wird ein Weg untersucht, den Konfigurationsraum des aktuellen Algorithmus zu reduzieren, indem eine alternative Herleitung herangezogen wird, die zur analytischen Berechnung gefalteter Integrale durch die Residuen einer inversen Laplace-Transformation führt.

Abstract

In plasma physics, models of warm dense matter like the uniform electron gas have become increasingly important. Because of the errors in some analytical approximations and the difficulties with some theoretical models, such as the simultaneous Coulomb interaction of all electrons in the uniform electron gas or the spin statistics which imply the Pauli principle, there is a need for *ab initio* simulations of the thermodynamic properties of the electron gas in different temperature ranges and coupling strengths.

One way to do this is with quantum Monte Carlo techniques which are based on Feynman's path integral formulation but suffer from the fermion sign problem. This can be partly circumvented by choosing different, complementary representations. One of these is given by Configuration Path Integral Monte Carlo, formulated in second quantization.

This bachelor's thesis provides two main ideas which could lead to improving Configuration Path Integral Monte Carlo. The first part explores higher order propagators in the derivation of the thermodynamic partition function and shows a canonical way to perform discrete simulations with arbitrary order. The second part shows a way to reduce the configuration space of the current algorithm by performing an alternative derivation which leads to the analytic computation of convoluted integrals by means of residues of an inverse Laplace transform.

Contents

Chapter 1

Introduction

Our world as we see it today is on small scales described by the incredibly accurate theory of quantum mechanics. Of all quantum mechanical particles, which are divided into fermi and bose particles, typical objects of interest are electrons and ions. Recently, an area of growing significance has been the description of systems of electrons under extreme conditions [1]. For example, the interior of giant planets and brown and white dwarfs is expected to contain warm dense matter [2, 3, 4], a state of high density and temperatures at the order of the Fermi energy. An important theoretical model as a description base is the uniform electron gas [5, 6]. Because of the great difficulties and approximations that have to be made in analytical calculations, it is of particular importance to be able to perform numerical simulations for computing thermodynamic properties.

One of the most successful computational methods at low temperatures is density functional theory [7, 8, 9]. However, especially for finite temperatures in the warm dense matter regime, these become unreliable so that other approaches are necessary. Very prominent are quantum Monte Carlo simulations, which are based on a path integral formulation [10]. These, though, are severely restricted due to the so-called fermion sign problem [11], which is shown to be NP-hard [12]. Since the fermion sign problem is representation-dependent, it can be partly circumvented by considering different bases which have different ranges of application [6, 13].

A representation which turned out to be complementary to the standard Path Integral Monte Carlo simulation performed in position space was Configuration Path Integral Monte Carlo, proposed by Schoof et al. [14, 15, 16]. Here, the operators were looked at in second quantization instead, while the sampled Monte Carlo paths resembled Fock states in occupation number representation. These two formulations, together with various other improvements and optimizations, cover a wide range of parameters, leaving only a gap at temperatures below the Fermi temperature $T < 0.5\,T_F$ with moderate coupling $r_s \approx 2 - 6$ [13].

This bachelor's thesis develops different approaches which may increase the range of applicability of the Configuration Path Integral Monte Carlo algorithm to greater coupling strength ($r_s \gtrsim 2$, depending on parameters like temperature and particle number) by trying to develop a second quantization analogue to Permutation Blocking Path Integral Monte Carlo by Dornheim et al [17, 18] and grouping configuration paths together. In this thesis we will consider two independent ideas.

After a brief introduction to the foundations of Monte Carlo and the Configuration Path Integral algorithm in particular, the first part addresses the possibility of higher order expansions, which naturally arise in the derivation of the path integral formulation of the canonical partition function. It is shown that higher order corrections to the partition function do not have a notable influence on the current Monte Carlo algorithm, which is based on so-called "continuous time". Simultaneously, the proof also provides a way to address Configuration Path Integral Monte Carlo in "discrete time", but with arbitrary order.

The second part of this thesis starts from the current formulation of the canonical partition function and aims to reduce the configuration space which is sampled during the Monte Carlo procedure, thus improving the algorithm. It is joined by a number of analytic aspects and ideas which provide several ways of implementing such a "reduced" algorithm and show starting points for alternative methods which could be investigated.

Chapter 2

Foundations

2.1 Second quantization and statistical physics

An essential quantity in statistical physics is the *canonical partition function* Z of a canonical ensemble, given by

$$Z = \operatorname{Tr} e^{-\beta H},$$

with the Hamiltonian H and inverse temperature $\beta = \frac{1}{k_B T}$. This can be used to derive the expectation values of certain observables of the system. As an example, its average energy is given by

$$\overline{E} = \langle H \rangle = -\frac{\partial}{\partial \beta} \ln(Z).$$

One way to evaluate the trace and simultaneously sample observable quantities is using Monte Carlo simulations. These depend on the chosen basis in which the trace will be calculated.

In *Configuration Path Integral Monte Carlo* (short: CPIMC) this is done using the second quantization occupation number representation of N-particle states. First, one should specify a complete orthonormal system (CONS) of underlying one-particles basis states $\{|s_i\rangle : i \in \mathbb{N}\}$ (e.g. plane waves). For our purposes we assume an ideal system in which the states of the N-particle Hilbert space $\mathcal{H}$ are tensor product states, i.e. $|s_{i_1} \cdots s_{i_N}\rangle = |s_{i_1}\rangle_1 \cdots |s_{i_N}\rangle_N$ and $\mathcal{H}^{(N)} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N$ with $|s\rangle_i \in \mathcal{H}_i$ where all the single-particle Hamiltonians h_i with Hilbert space $\mathcal{H}_i$ have the same structure.

In a physical system of identical particles this representation carries redundant information as there can be no difference between the physical states of e.g. $|s_1 s_2\rangle$ and $|s_2 s_1\rangle$. In other words, we cannot know whether particle 1 is in state 1 and particle 2 is in state 2 or it is the other way around. For the fermionic systems of our interest exchanging any two particles must only change the phase of the state by π (or equivalently, flip the sign): $|s_1 s_2\rangle = -|s_2 s_1\rangle$ This enforces the Pauli principle so that no state can be occupied twice.

Note that this is not in contradiction to the indistinguishability of the corresponding physical states. Any measurable physical quantity will depend only on the absolute value of the amplitude squared so that the signs cancel.

Antisymmetrizing the states removes the unphysical redundancy:

$$\begin{aligned}|s_{i_1}\cdots s_{i_N}\rangle_- &:= \frac{1}{\sqrt{N!}}\sum_{\pi\in S_N}\operatorname{sgn}(\pi)|s_{i_1}\rangle_{\pi(1)}\cdots|s_{i_N}\rangle_{\pi(N)}\\ &= \frac{1}{\sqrt{N!}}\det\left(|s_{i_k}\rangle_l\right)_{k,l=1}^{N}.\end{aligned}$$

In occupation number representation we now first agree on an (arbitrary but fixed) order of the single-particle basis states. Since we have already established a numbering, we will order the basis states accordingly by $|s_1\rangle, |s_2\rangle, \ldots$.

We now write this as

$$|n\rangle := |n_1\, n_2\, n_3\cdots\rangle$$

where $n_i \in \{0,1\}$ denotes the number of particles in basis state $|s_i\rangle$. As an example, we would have $|s_3 s_1\rangle_- = -|s_1 s_3\rangle_- = (-1)\cdot|1\,0\,1\,0\,0\,0\cdots\rangle$. These properly antisymmetrized N-particle states $\{|n\rangle : n \in \{0,1\}^{\mathbb{N}}\}$ inherit from $\{|s_i\rangle : i \in \mathbb{N}\}$ the property of constituting a complete orthonormal system for the antisymmetrized N-particle Hilbert space $\mathcal{H}_-^{(N)}$.

The formulation established up to this point might already seem sufficient. One key ingredient though that is still missing is the introduction of creation and annihilation operators, which will allow us to express previously known operators in the second quantization formalism.

For this we first allow changes in the particle number by looking at the exterior algebra, also called *Fock space*,

$$\mathcal{F} := \bigoplus_{N=0}^{\infty}\mathcal{H}_-^{(N)} = \mathcal{H}_-^{(0)}\oplus\mathcal{H}_-^{(1)}\oplus\mathcal{H}_-^{(2)}\oplus\ldots.$$

Next, define the *creation operator* $a_i^\dagger$ and the *annihilation operator* a_i (which are mutually adjoint) by

$$\begin{aligned}a_i^\dagger\,|\cdots n_i\cdots\rangle &:= (-1)^{\sum_{k=0}^{i-1}n_k}\,\delta_{n_i,0}|\cdots n_i+1\cdots\rangle\\ a_i\,|\cdots n_i\cdots\rangle &:= (-1)^{\sum_{k=0}^{i-1}n_k}\,\delta_{n_i,1}|\cdots n_i-1\cdots\rangle.\end{aligned}$$

The sign is chosen is such a way that for any $i,j\in\mathbb{N}$ the commutation relations

$$\{a_i^\dagger, a_j^\dagger\} = 0 = \{a_i, a_j\},\quad \{a_i^\dagger, a_j\} = \delta_{ij}$$

are satisfied, where the *anti-commutator* has been defined as $\{A, B\} := AB + BA$. As a side result we note that these operators provide a way of creating all Fock states from the *vacuum state* $|0\rangle := |0\,0\,0\cdots\rangle$. As an example, $|101000\cdots\rangle = a_1^\dagger a_3^\dagger |0\rangle$. More generally,

$$|n\rangle = \left(\prod_{i=1}^{\infty} \left(a_i^\dagger\right)^{n_i}\right) |0\rangle,$$

where one has to be careful with preserving the order.

An important theorem is that any one-particle operator $B = \sum_{i=1}^{\infty} B_i$ (B_i acts on the i-th particle) can be expressed in second quantization as [19]

$$B = \sum_{i,j=1}^{\infty} b_{ij} a_i^\dagger a_j$$

with $b_{ij} := \langle s_i|B|s_j\rangle$. Similarly, any two-particle operator $B = \frac{1}{2}\sum_{i,j=1}^{\infty} B_{i,j}$ can be expressed as

$$B = \frac{1}{2} \sum_{i,j,k,l=1}^{\infty} b_{ijkl} a_i^\dagger a_j^\dagger a_l a_k.$$

2.2 Monte Carlo simulation

2.2.1 Configuration Path Integral Monte Carlo

Suppose we have an N-particle system and want to calculate $Z = \operatorname{Tr} e^{-\beta H}$ in second quantization. A simple idea would be to define $\varepsilon := \beta/M$ (with $M \in \mathbb{N}$) and evaluate

$$\begin{aligned} Z = \operatorname{Tr} e^{-\beta H} &= \sum_n \langle n|e^{-\beta H}|n\rangle = \sum_n \langle n| \left(e^{-\varepsilon H}\right)^M |n\rangle \\ &\approx \sum_n \sum_{n_1} \cdots \sum_{n_{M-1}} \langle n|1-\varepsilon H|n_1\rangle \cdots \langle n_{M-1}|1-\varepsilon H|n\rangle. \end{aligned}$$

The summations are over all N-particle Fock basis states $|n\rangle \in \mathcal{H}_-^{(N)}$, which we have mentioned earlier to constitute a complete orthonormal basis for $\mathcal{H}_-^{(N)}$. This idea lies at the heart of Configuration Path Integral Monte Carlo, whose de facto representation of Z is exhibited in greater detail in Chapter 3.

The use of Monte Carlo techniques now stems from the fact that there is a large number of possible configurations involved which have to be summed up (later, some integrals will be involved as well). The high dimension of this summation / integration practically prohibits most other techniques.

2.2.2 Metropolis algorithm

In principle, we arrive (by CPIMC or some other formulation) at an expression of the form

$$Z = \sum\!\!\!\!\!\!\int_{\mathcal{C}} W(\mathcal{C}),$$

where the summation / integration is executed over all possible paths $\mathcal{C}$ with corresponding weights $W(\mathcal{C})$. In the CPIMC example from above, a path would be determined by $\mathcal{C} = (n, n_1, \ldots, n_M)$ with given M, whereas the actual CPIMC formulation derived in Chapter 3 has configurations $\mathcal{C} = (K, n, n_1, \ldots, n_{K-1}, \tau_1, \ldots, \tau_K)$ with the number of kinks (configuration changes) K and imaginary times τ_i at which these changes occur.

Now, our goal is to calculate expectation values of observables A given in the form

$$\langle A \rangle = \frac{1}{Z} \sum\!\!\!\!\!\!\int_{\mathcal{C}} A(\mathcal{C}) W(\mathcal{C}) = \frac{\sum\!\!\!\!\!\!\int_{\mathcal{C}} A(\mathcal{C}) W(\mathcal{C})}{\sum\!\!\!\!\!\!\int_{\mathcal{C}} W(\mathcal{C})}.$$

This could be interpreted as sampling A by a probability distribution $W(\mathcal{C})$ if all $W(\mathcal{C})$ were positive. One can always achieve this by defining the sign $S(\mathcal{C}) := \mathrm{sgn}(W(\mathcal{C}))$ and writing

$$\langle A \rangle = \frac{\sum\!\!\!\!\!\!\int_{\mathcal{C}} A(\mathcal{C}) S(\mathcal{C}) |W(\mathcal{C})|}{\sum\!\!\!\!\!\!\int_{\mathcal{C}} S(\mathcal{C}) |W(\mathcal{C})|} = \frac{\langle AS \rangle'}{\langle S \rangle'}$$

with the ensemble averaging $\langle \cdot \rangle'$ via $Z' := \sum\!\!\!\!\!\!\int_{\mathcal{C}} |W(\mathcal{C})|$. This is where the *fermion sign problem* arises, as small average signs $\langle S \rangle'$ lead to large uncertainty in measuring $\langle A \rangle$.

Another problem is our ignorance of the value of Z', which is needed to ensure the weights are properly normalized to $p(\mathcal{C}) := |W(\mathcal{C})|/Z'$ and really resemble a probability distribution. In the *Metropolis algorithm* [20] this is circumvented by only looking at transition probabilities $p(\mathcal{C}_i \to \mathcal{C}_j)$ and requiring they meet the *detailed balance condition*

$$p(\mathcal{C}_i)\, p(\mathcal{C}_i \to \mathcal{C}_j) = p(\mathcal{C}_j)\, p(\mathcal{C}_j \to \mathcal{C}_i).$$

One can verify that a solution is given by

$$p(\mathcal{C}_i \to \mathcal{C}_j) = \min\left\{1, \frac{p(\mathcal{C}_j)}{p(\mathcal{C}_i)}\right\} = \min\left\{1, \frac{|W(\mathcal{C}_j)|}{|W(\mathcal{C}_i)|}\right\},$$

which removes the need to know the normalization factor Z'.

Finally, and roughly speaking, one can perform the simulation by generating a *Markov chain*, i.e. starting with configuration $\mathcal{C}_0$ and proposing $\mathcal{C}_{i+1}$ with the help of $\mathcal{C}_i$ and $p(\mathcal{C}_i \to \mathcal{C}_{i+1})$. The resulting states $\{\mathcal{C}_i : i = 1, 2, \ldots\}$ are known to be distributed according to $p(\mathcal{C})$. For more details see e.g. [21, 16].

Part I

Higher Order CPIMC

Chapter 3

Canonical Partition Function

We now begin with deriving the standard continuous time formulation of the partition function in CPIMC [14, 15]. One way to do this is starting with a discrete time formulation and performing the limit to continuous time. We choose this path as it provides a way to investigate higher order corrections. Unless otherwisely stated, we will always identify $n_0 \equiv n \equiv n_K$, $\tau_0 \equiv 0$ and $\tau_{K+1} \equiv \beta$.

The usual problem is that, in $Z = \operatorname{Tr} e^{-\beta H}$, the offdiagonal part of the Hamiltonian H prohibits an exact evaluation. In the following we will split H into a diagonal and an offdiagonal part, writing $H = D + Y$. This means that D and Y are constructed in such a way that $\langle n|D|m\rangle \sim D_n \, \delta_{nm}$ and $\langle n|Y|n\rangle = 0$. For $M \in \mathbb{N}$ define $\varepsilon := \frac{\beta}{M}$. We can then write

$$Z = \operatorname{Tr} e^{-\beta H} = \sum_n \langle n|e^{-\beta H}|n\rangle = \sum_n \langle n| \left(e^{-\varepsilon H}\right)^M |n\rangle,$$

where the summation is executed over every N-particle Fock state. This enables us to approximate

$$e^{-\varepsilon H} = 1 - \varepsilon H + \mathcal{O}(\varepsilon^2),$$

up to second order. Taking the product of M such approximations therefore yields

$$\begin{aligned} Z &= \sum_n \langle n| \left(1 - \varepsilon H + \mathcal{O}(\varepsilon^2)\right)^M |n\rangle = \sum_n \langle n|(1 - \varepsilon H)^M|n\rangle + \mathcal{O}(\varepsilon) \\ &= \sum_{n_0 = n_M} \sum_{n_1} \cdots \sum_{n_{M-1}} \prod_{i=1}^{M} \langle n_{i-1}|1 - \varepsilon H|n_i\rangle + \mathcal{O}(\varepsilon). \end{aligned}$$

If we now split H into D and Y, we get

$$\langle n|1 - \varepsilon(D + Y)|m\rangle = \begin{cases} 1 - \varepsilon D_n & n = m \\ -\varepsilon Y_{n,m} & n \neq m \end{cases}.$$

Here we have defined $D_n := \langle n|D|n\rangle$ and $Y_{n,m} := \langle n|Y|m\rangle$. One can also write

$$\langle n|1-\varepsilon H|m\rangle = \begin{cases} e^{-\varepsilon D_n} & n=m \\ -\varepsilon Y_{n,m} & n\neq m \end{cases} + \mathcal{O}(\varepsilon^2)$$

instead, committing an error of magnitude $\mathcal{O}(\varepsilon^2)$, which does not change the total order of the partition sum. Hence, rewriting the summation of the partition sum based on the latter expression and in terms of the number of kinks (a *kink* is defined as a Y-term in the summation. Physically, this is where the configurations change in a given path), we get

$$Z = \sum_{\substack{K=0\\K\neq 1}}^{M} \sum_{l_1=1}^{M} \cdots \sum_{l_K=l_{K-1}+1}^{M} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} e^{-\varepsilon \sum_{i=0}^{K} D_{n_i}(l_{i+1}-l_i-1)} (-\varepsilon)^K \prod_{i=1}^{K} Y_{n_{i-1},n_i}.$$

Here we implicitly require $n_i \neq n_{i+1}$ for all i. Also, $n_0 := n =: n_K$, $l_0 := 0$ and $l_{K+1} := M+1$ in the summation. In the continuous time limit $M\to\infty$ we have

$$\varepsilon l_i \to \tau_i \text{ and } \sum_{l_i=l_{i-1}+1}^{M} \varepsilon \to \int_{\tau_{i-1}}^{\beta} d\tau_i.$$

Therefore we arrive at

$$\begin{aligned} Z &= \sum_{K=0}^{\infty} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} \int_0^\beta d\tau_1 \cdots \int_{\tau_{K-1}}^{\beta} d\tau_K \; (-1)^K \, e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} Y_{n_{i-1},n_i} \qquad (3.1) \\ &\equiv \sum\!\!\!\!\!\!\int_{\mathcal{C}} W(\mathcal{C}), \end{aligned}$$

where a path $\mathcal{C}$ is determined by $\mathcal{C} = (K, n_0, \ldots, n_{K-1}, \tau_1, \ldots, \tau_K)$ and has a weight

$$W(\mathcal{C}) = (-1)^K \, e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} Y_{n_{i-1},n_i}.$$

This formulation has proven itself, especially for the uniform electron gas at warm dense matter conditions, numerous times [14, 22, 13]. In this model the Hamiltonian consists of Coulomb-interacting electrons with a compensating positive background charge. The two configuration parameters of the uniform electron gas are the reduced temperature $\theta = k_B T/E_F$ (with the Fermi energy E_F) and the Wigner-Seitz radius

$$r_s = \left(\frac{3}{4\pi n}\right)^{1/3} \cdot \frac{1}{a_B},$$

in units of the Bohr radius a_B and dependent on the density n. Due to the fermion sign problem, CPIMC is restricted to $\theta/r_s \gtrsim 1$ and applicable up to densities of order $r_s \sim 1$.

Chapter 4

Higher Order Approximation

We will now generalize the previous expression by introducing factors c_i, $i \in \{1, \dots, M\}$. The idea is the following: If we replace expressions of the type $(1-\varepsilon H)$ by $(1-c_i\varepsilon H)$, the c_i could maybe be optimized to yield an overall higher order of the discrete-time partition function

$$Z = \sum_n \langle n|(1 - c_1\varepsilon H)\cdots(1 - c_M\varepsilon H)|n\rangle.$$

Since M becomes arbitrarily large, optimizing all the c_i independently is a difficult task. We will therefore handle the c_i by introducing a periodicity, i.e. $c_i = c_{i+P}$ for some $P \in \mathbb{N}$. Then, taking the limit $P \to \infty$ becomes the final expression.

The idea is to try to find a factorization which results in a trace of higher order, similar to the work of Chin and Sakkos [23, 24] which led to Permutation Blocking PIMC [17, 18]. A problem is that until now there is no known way of evaluating $e^{-\varepsilon Y}$ (Y is the offdiagonal part of the Coulomb potential) in second quantization. Previously, this was always approximated linearly. The following sections seek to improve the order via optimizing the c_i, without having to evaluate higher order terms like Y^2 or Y^3 directly.

4.1 Periodicity P = 2

Since M is very large, we can always assume that $M \in P\ \mathbb{N} = \{P, 2P, 3P, \dots\}$. In the limit $M \to \infty$ this does not make a difference. In the following, let $P = 2$. One can write

$$e^{-\beta H} = \left(e^{-2\varepsilon H}\right)^{M/2}.$$

In the standard expression we would then proceed to approximate

$$e^{-2\varepsilon H} = e^{-\varepsilon H}e^{-\varepsilon H} = (1 - \varepsilon H)(1 - \varepsilon H) + \mathcal{O}(\varepsilon^2).$$

We will now try to increase this order of approximation by introducing c_1 and c_2. First, Taylor-expanding up to $\mathcal{O}(\varepsilon^3)$, we have

$$e^{-2\varepsilon H} = 1 - 2\varepsilon H + 2(\varepsilon H)^2 + \mathcal{O}(\varepsilon^3).$$

If we could now find some c_i such that

$$1 - 2\varepsilon H + 2(\varepsilon H)^2 = (1 - c_1\varepsilon H)(1 - c_2\varepsilon H),$$

we are finished and have

$$e^{-2\varepsilon H} = (1 - c_1\varepsilon H)(1 - c_2\varepsilon H) + \mathcal{O}(\varepsilon^3).$$

The existence of such c_i is given by the following lemma.

Lemma 4.1. *For every $P \in \mathbb{N}$ there exist $c_1, \dots, c_P \in \mathbb{C} \setminus \{0\}$ such that*

$$e^{-P\varepsilon H} = (1 - c_1\varepsilon H) \cdots (1 - c_P\varepsilon H) + \mathcal{O}(\varepsilon^{P+1}).$$

Proof. Let $P \in \mathbb{N}$. Define the formal polynomial

$$f_P(X) := \sum_{n=0}^{P} \frac{(-P)^n}{n!} X^n \equiv \sum_{n=0}^{P} a_n X^n \in \mathbb{R}[X].$$

It is evident that $e^{-P\varepsilon H} = f_P(\varepsilon H) + \mathcal{O}(\varepsilon^{P+1})$. By the fundamental theorem of algebra the complex numbers are algebraically closed. Equivalently, there exist $x_1, \dots, x_P \in \mathbb{C}$, the roots of f_P, such that

$$f_P(X) = \frac{P^P}{P!}(x_1 - X) \cdots (x_P - X).$$

By comparing coefficients we get

$$1 = a_0 = \frac{P^P}{P!} x_1 \cdots x_P.$$

This also implies that all $x_i \neq 0$. Factoring out results in

$$f_P(X) = \underbrace{\frac{P^P}{P!} x_1 \cdots x_P}_{=1} \left(1 - \frac{1}{x_1}X\right) \cdots \left(1 - \frac{1}{x_P}X\right).$$

Therefore, defining $c_i := 1/x_i \neq 0$, we have $f_P(X) = (1 - c_1X) \cdots (1 - c_PX)$ and thus

$$e^{-P\varepsilon H} = f_P(\varepsilon H) + \mathcal{O}(\varepsilon^{P+1}) = (1 - c_1\varepsilon H) \cdots (1 - c_P\varepsilon H) + \mathcal{O}(\varepsilon^{P+1}).$$

□

We do not have to know the actual values of the c_i, though for $P = 2$ it is straightforwardly checked that $c_1 = 1 + i$ and $c_2 = 1 - i$ satisfy the requirements. This creates another problem: Because the c_i are in general not real but complex, we run into trouble with the Monte Carlo simulation. This can be fixed by regrouping terms. First, though, we will have a closer look at the c_i.

By comparing coefficients in the proof of the previous lemma we get various useful identities. As an example, in the case $P = 2$, we have $1-2X+2X^2 = (1-c_1X)(1-c_2X) = 1 - (c_1 + c_2)X + c_1c_2X^2$. This means $c_1 + c_2 = 2$ and $c_1c_2 = 2$, which we will use in the following. Similar identities hold for $P > 2$.

The terms are now regrouped depending on the so-called "kink order". In the case of $P = 2$ we differentiate between no kinks, simple kinks and twofold kinks (This does not refer to the kink type, i.e. $T2$ or $T4$, which arise with the Coulomb potential in second quantization [15]). The goal is to evaluate

$$\langle n|e^{-2\varepsilon H}|m\rangle = \sum_{n_1}\langle n|1 - c_1\varepsilon H|n_1\rangle\langle n_1|1 - c_2\varepsilon H|m\rangle + \mathcal{O}(\varepsilon^3).$$

Case 1: No kink. This is only the case if $n = n_1 = m$. Then we get

$$\begin{aligned}\langle n|1 - c_1\varepsilon H|n_1\rangle\langle n_1|1 - c_2\varepsilon H|m\rangle &= (1 - c_1\varepsilon D_n)(1 - c_2\varepsilon D_n)\\ &= f_2(\varepsilon D_n) = e^{-2\varepsilon D_n} + \mathcal{O}(\varepsilon^3).\end{aligned}$$

Case 2: Simple kink. If $n = m$, we would either have zero or two kinks. Therefore, in this case we must have $n \neq m$. Also, $n_1 = n$ or $n_1 = m$. Else there would be two kinks as well. If $n_1 = n$, then

$$\langle n|1 - c_1\varepsilon H|n\rangle\langle n|1 - c_2\varepsilon H|m\rangle = (1 - c_1\varepsilon D_n)(-c_2\varepsilon Y_{n,m}).$$

Similarly, for $n_1 = m$

$$\langle n|1 - c_1\varepsilon H|m\rangle\langle m|1 - c_2\varepsilon H|m\rangle = (-c_1\varepsilon Y_{n,m})(1 - c_2\varepsilon D_m).$$

Since both cases occur in the partition function, we can add them together to get

$$\begin{aligned}-\varepsilon Y_{n,m}(c_2 - c_1c_2\varepsilon D_n + c_1 - c_1c_2\varepsilon D_m) &= -\varepsilon Y_{n,m}(c_1 + c_2 - c_1c_2\varepsilon(D_n + D_m))\\ &= -\varepsilon Y_{n,m}(2 - 2\varepsilon(D_n + D_m))\\ &= -2\varepsilon\ Y_{n,m}e^{-\varepsilon(D_n+D_m)} + \mathcal{O}(\varepsilon^3)\end{aligned}$$

Grouping the terms in this way therefore results in real-valued weights.

Case 3: Twofold kink. For a twofold kink we have $n \neq n_1 \neq m$ (this does **not** imply $n \neq m$. $n = m$ is still possible). Therefore, with $c_1 c_2 = 2$,

$$\langle n|1 - c_1\varepsilon H|n_1\rangle\langle n_1|1 - c_2\varepsilon H|m\rangle = (-c_1\varepsilon Y_{n,n_1})(-c_2\varepsilon Y_{n_1,m}) = (-2\varepsilon)^2 \frac{1}{2!} Y_{n,n_1} Y_{n_1,m}.$$

With this, the canonical partition function can be written in a similar fashion as before. First, though, we introduce some new terminology.

K. K will always depict the total number of kinks, counted with their order (a twofold kink adds two kinks to the total number of kinks).

N_p. For given kink positions $l_1, \ldots, l_K$, $N_p = N_p(l_1, \ldots, l_K)$ will denote the number of p-fold kinks. We therefore have $\sum_{p=1}^{\infty} p \cdot N_p = K$.

As an example: Let $(l_1, \ldots, l_7) = (1, 2, 2, 2, 4, 5, 9)$. Then $N_1 = 4$, $N_3 = 1$ and $N_p = 0$ for $p \notin \{1, 3\}$: We have "three kinks at once" at discrete time position 2 and four simple kinks at times $1, 4, 5$ and 9.

R. Again for $l_1, \ldots, l_K$ given, $R = R(l_1, \ldots, l_k)$ is a kink-dependent "remainder" term given by

$$R = \prod_{p=1}^{\infty} \left(\frac{1}{p!}\right)^{N_p}.$$

Since almost all N_p are zero, R does not diverge for any configuration $l_1, \ldots, l_K$. In the example from above we would find $R = 1^4 \cdot (\frac{1}{3!})^1 = \frac{1}{6}$.

$\delta_{N \leq p}$. This Kronecker delta is 1 if for all $q > p$ $N_q = 0$, i.e., $\delta_{N \leq p} := \prod_{q=p+1}^{\infty} \delta_{N_q,0}$.

T_i. For given $l_1, \ldots, l_K$ and $i \in \{1, \ldots, K\}$, T_i denotes the order of the kink at location l_i. That is, $T_i = T_i(l_1, \ldots, l_K) = |\{j \in \{1, \ldots, K\} : l_j = l_i\}| \geq 1$. In the previous example this amounts to $T_1 = 1$, $T_2 = 3$, $T_3 = 0$, $T_4 = 1$ and so on.

We can now write the canonical partition function as

$$Z = \sum_{\substack{K=0 \\ K \neq 1}}^{M} \sum_{l_1=1}^{M/2} \cdots \sum_{l_K=l_{K-1}}^{M/2} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} e^{-\varepsilon \sum_{i=0}^{K} \left(D_{n_i} 2(l_{i+1} - l_i - 1) + r_i\right)} (-2\varepsilon)^K \prod_{i=1}^{K} Y_{n_{i-1},n_i} \; R\delta_{N \leq 2}.$$

Here we have defined (with $T_{K+1} := 2$) an extra factor depending on the kink order,

$$r_i := \begin{cases} D_{n_i} + D_{n_{i+1}} & T_{i+1} = 1 \\ 0 & T_{i+1} = 2 \end{cases}.$$

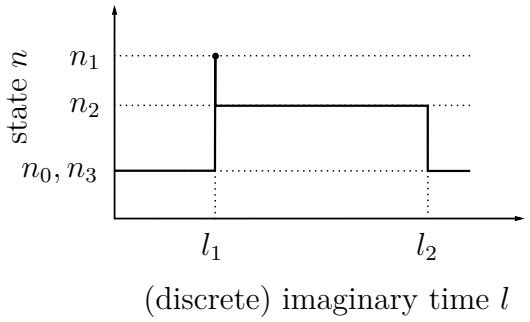

Figure 4.1: At (discrete) imaginary time l_1 is a second order kink: The path goes from n_0 over n_1 to n_2, all in one instant. At l_2 is a first order kink from n_2 to $n_3 = n_0$.

Note the resemblance to the first order canonical partition function in the previous chapter. We can actually write the former version similarly to the latter: Change the summation from $l_i = l_{i-1} + 1$ to $l_i = l_{i-1}$ and compensate this by adding $\delta_{N\leq 1}$. Next, add R (as this will always be equal to one). Note that all the occurrences of 2 are due to the periodicity $P = 2$. In the previous formulation [14], which corresponds to $P = 1$, they disappear. Finally, define $r_i := 0$.

The limit $M \to \infty$. We perform the limit to continuous time in a similar fashion as before. The difference is that we now have

$$2\varepsilon\, l_i \to \tau_i \text{ and } \sum_{l_i = l_{i-1}}^{M/2} 2\varepsilon \to \int_{\tau_{i-1}}^{\beta} d\tau_i.$$

First, observe that the $-\varepsilon\, r_i$ terms in the exponent vanish with $\varepsilon \to 0$. We then arrive at

$$Z = \sum_{\substack{K=0 \\ K\neq 1}}^{M} \int_0^\beta d\tau_1 \cdots \int_{\tau_{K-1}}^\beta d\tau_K \sum_n \sum_{n_1} \cdots \sum_{n_{K-1}} e^{-\sum_{i=0}^K D_{n_i}(\tau_{i+1}-\tau_i)} (-1)^K \prod_{i=1}^K Y_{n_{i-1},n_i}\, R\, \delta_{N\leq 2}.$$

This differs from the result of the previous chapter (corresponding to $P = 1$) only by the factor $R\delta_{N\leq 2}$. First, this does not render the previous $P = 1$ case wrong: We have $R = 1$ almost everywhere. Only when $\tau_i = \tau_{i+1}$ for some i we get a different weight. But this is a set of measure zero and therefore does not change the total integration result Z.

Second, let us look at the $P = 1$ case again: $\delta_{N\leq 1}$ ensures that R is always equal to one since there are only simple kinks allowed. Thus, the missing of this factor was not noticeable previously. Still, the second order result suggests that one has to prohibit $\tau_i = \tau_{i+1}$ in the classical algorithm or allow it but compensate this by adding the corresponding factor R.

Third, we will generalize this to an arbitrary periodicity P in the following section.

As the above expression for Z suggests, we will find that in the continuous limit the only difference is a factor $R\delta_{N\leq P}$. Taking the limit $P \to \infty$ yields $\delta_{N\leq P} \to 1$, so merely the additional factor $R = R(\tau_1, \ldots, \tau_K) = \prod_{p=1}^{\infty} \left(\frac{1}{p!}\right)^{N_p(\tau_1,\ldots,\tau_K)}$ remains. We arrive at

$$Z \equiv \sum\!\!\!\!\!\!\int_{\mathcal{C}} W(\mathcal{C})$$

with paths $\mathcal{C}$ still determined by $\mathcal{C} = (K, \tau_1, \ldots, \tau_k, n_0, \ldots, n_{K-1})$ and

$$W(\mathcal{C}) = e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1}-\tau_i)}(-1)^K \prod_{i=1}^{K} Y_{n_{i-1},n_i}\ R(\tau_1, \ldots, \tau_K).$$

4.2 Arbitrary periodicity

Now assume arbitrary periodicity P and, again, $M \in P\,\mathbb{N}$. We write

$$e^{-\beta H} = (e^{-P\varepsilon H})^{M/P}$$

with $\varepsilon = \beta/M$. By Lemma 4.1 we find $c_1, \ldots, c_P \neq 0$ satisfying

$$e^{-P\varepsilon H} = (1 - c_1\varepsilon H)\cdots(1 - c_P\varepsilon H) + \mathcal{O}(\varepsilon^{P+1}).$$

As in the previous section, we insert identities to get

$$\langle n|e^{-P\varepsilon H}|n_P\rangle = \sum_{n_1}\cdots\sum_{n_{P-1}} \langle n|1 - c_1\varepsilon H|n_1\rangle \ldots \langle n_{P-1}|1 - c_P\varepsilon H|n_P\rangle + \mathcal{O}(\varepsilon^{P+1}).$$

Again we can now have kink orders from zero up to P. Let $\kappa \in \{0, \ldots, P\}$ be the kink order. For given kink order κ and starting position n, this means that there must be κ distinct states $\in \{|n_1\rangle, \ldots, |n_P\rangle\}$ which differ from their predecessors. Each of those higher order kinks should be characterized by κ and the $\kappa + 1$ states $\{|n^{(0)}\rangle = |n\rangle, |n^{(1)}\rangle, \ldots, |n^{(\kappa)}\rangle\}$ We therefore follow the lead of case $P = 2$ and group all possible summation entries together which belong to the given higher order kink:

Case 1: No kink. In this case we have only one state $|n\rangle$ and therefore

$$\langle n|1 - c_1\varepsilon H|n\rangle \cdots \langle n|1 - c_P\varepsilon H|n\rangle = (1 - c_1\varepsilon D_n)\cdots(1 - c_P\varepsilon D_n) = e^{-P\varepsilon D_n} + \mathcal{O}(\varepsilon^{P+1})$$

by definition of the c_i.

Case 2: Kink of order $\kappa \geq 1$. Here we have states $|n\rangle, \ldots, |n^{(\kappa)}\rangle$. This implies a non-empty summation and a term of

$$\sum_{1\leq l_1<\cdots<l_\kappa\leq P} \prod_{i_0=1}^{l_1-1} \langle n^{(0)}|1-c_{i_0}\varepsilon H|n^{(0)}\rangle \cdot \langle n^{(0)}|1-c_{l_1}\varepsilon H|n^{(1)}\rangle$$
$$\cdot \ldots \cdot \prod_{i_{\kappa-1}=l_{\kappa-1}+1}^{l_\kappa-1} \langle n^{(\kappa-1)}|1-c_{i_{\kappa-1}}\varepsilon H|n^{(\kappa-1)}\rangle \cdot \langle n^{(\kappa-1)}|1-c_{l_{\kappa-1}}\varepsilon H|n^{(\kappa)}\rangle$$
$$\cdot \prod_{i_\kappa=l_\kappa+1}^{P} \langle n^{(\kappa)}|1-c_{i_\kappa}\varepsilon H|n^{(\kappa)}\rangle.$$

This also starts to resemble a path in the continuous time limit: At positions l_i we have a jump, else the configurations stay the same. We factor out the offdiagonal terms and write this as

$$(-\varepsilon)^\kappa Y_{n^{(0)},n^{(1)}} \cdots Y_{n^{(\kappa-1)},n^{(\kappa)}} \sum_{1\leq l_1<\cdots<l_\kappa\leq P} \prod_{i_0=1}^{l_1-1} \langle n^{(0)}|1-c_{i_0}\varepsilon H|n^{(0)}\rangle$$
$$\cdot \ldots \cdot \prod_{i_\kappa=l_\kappa+1}^{P} \langle n^{(\kappa)}|1-c_{i_\kappa}\varepsilon H|n^{(\kappa)}\rangle \cdot c_{l_1}\cdots c_{l_\kappa}.$$

The summation term, when substituting $\langle n^{(i)}|1-c_j\varepsilon H|n^{(i)}\rangle = 1-c_j\varepsilon D_{n^{(i)}}$, is just a polynomial in ε. Luckily this polynomial does not have to be known exactly. Its order 0 term though is given by

$$\sum_{1\leq l_1<\cdots<l_\kappa\leq P} c_{l_1}\cdots c_{l_{\kappa-1}} = \frac{P^\kappa}{\kappa!}.$$

This can be shown by looking at the proof of Lemma 4.1. Clearly, by its last expression for f_P, the above is just equal to the term of order κ of the polynomial $f_P(-X)$. But this is just, by definition of f_P, $\frac{P^\kappa}{\kappa!}$.

Factoring this out results in

$$\frac{(-P\varepsilon)^\kappa}{\kappa!} Y_{n^{(0)},n^{(1)}} \cdots Y_{n^{(\kappa-1)},n^{(\kappa)}} \cdot g(\varepsilon)$$

with a polynomial g of the form $g(X) = 1 + a_1 X + \cdots + a_{P-\kappa}X^{P-\kappa}$. The exact expression for g does not matter. In fact, in the previous case of $P = 2$ replacing it by an exponential (in such a way that the error is of magnitude $\mathcal{O}(\varepsilon^{\kappa+1})$) was merely a cosmetic modification. This is where the $e^{-\varepsilon r_i}$ extra factor came from, which vanishes in the continuous time limit. Next, perform said limit. We can write Z as

$$Z = \sum_{\substack{K=0\\K\neq 1}}^{M} \sum_{l_1=1}^{M/P} \cdots \sum_{l_K=l_{K-1}}^{M/P} \sum_n \sum_{n_1} \cdots \sum_{n_{K-1}} e^{-P\varepsilon\sum_{i=0}^K D_{n_i}(l_{i+1}-l_i-1)} (-P\varepsilon)^K g(\varepsilon) \prod_{i=1}^{K} Y_{n_{i-1},n_i}\, R\, \delta_{N\leq P}.$$

As before, the sum limit of M/P comes from the fact that we factorized

$$e^{-\beta H} = (e^{-P\varepsilon H})^{M/P}.$$

In the continuous time limit $M \to \infty$ this becomes

$$P\varepsilon l_i \to t_i \text{ and } \sum_{l_i=l_{i-1}}^{M/P} P\varepsilon \to \int_{\tau_{i-1}}^{\beta} d\tau_i.$$

Also, because $\varepsilon = \beta/M \to 0$, $g(\varepsilon) \to 1$ (this is what was mentioned earlier: The extra factor vanishes). The remaining expression is

$$Z = \sum_{\substack{K=0 \\ K\neq 1}}^{M} \int_0^{\beta} d\tau_1 \cdots \int_{\tau_{K-1}}^{\beta} d\tau_K \sum_n \sum_{n_1} \cdots \sum_{n_{K-1}} e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1}-\tau_i)} (-1)^K \prod_{i=1}^{K} Y_{n_{i-1},n_i} \, R \, \delta_{N\leq P}.$$

We have therefore proved

Theorem 4.2. *In the continuous time limit, higher order corrections to the partition function Z occur only on a set of measure zero.*

Any factorization of the exponential $e^{-\beta H}$ will necessarily have to coincide with $e^{-\beta H}$ itself up to some order.

Example 4.3. First, as always, denote $\varepsilon := \beta/M$ and $H = D + Y$. Now assume we would like to investigate a higher order expansion of $e^{-\varepsilon H}$ in the form of $e^{-\varepsilon H} = 1 - \varepsilon H + \frac{\varepsilon^2}{2} H^2 + \mathcal{O}(\varepsilon^3) = 1 - \varepsilon D - \varepsilon Y + \frac{\varepsilon^2}{2} D^2 + \frac{\varepsilon^2}{2} Y^2 + \frac{\varepsilon^2}{2} DY + \frac{\varepsilon^2}{2} YD + \mathcal{O}(\varepsilon^3)$. By the previous theorem, or explicitly by the section on periodicity $P = 2$ (replacing $\varepsilon' := \varepsilon/2$, $M' := 2M$), this would yield in the continuous time limit only a minor change on a set of measure zero, which does not have an impact on Z.

Example 4.4. In the next example, again with $\varepsilon := \beta/M$ and $H = D + Y$, we express $e^{-\varepsilon H} = e^{-\frac{\varepsilon}{2} D} e^{-\varepsilon Y} e^{-\frac{\varepsilon}{2} D} + \mathcal{O}(\varepsilon^3)$. Assume we would like to preserve the property that this is accurate up to $\mathcal{O}(\varepsilon^3)$. Still, this differs to $1 - \varepsilon H + \frac{\varepsilon^2}{2} H^2$ only by an error of $\mathcal{O}(\varepsilon^3)$ and therefore yields the same partition function as the previous example in the continuous time limit.

Concluding, we have shown that in the continuous time limit higher order corrections, though they do exist, do not impact Z because they only occur on configurations in a set of measure zero. This limits higher order expansions to discrete time CPIMC, for which we have given in theory an algorithm for evaluating an arbitrary expansion order by first factorizing using complex numbers and then regrouping these to stay at real-valued weights.

Part II

Simple CPIMC

Chapter 5

Simplified Partition Function

5.1 Derivation of a simplified partition function

In the following, we will derive another formulation of the Configuration Path Integral Monte Carlo algorithm. Because of the high level of complexity in some formulas, Section 7.3 provides a fully fledged example to follow along with. It will therefore occasionally be referenced to, when expedient.

The reformulation can be started in three different ways. One of them is a complex factorization as before and the second amounts to evaluating the integrals in Eq. (3.1). The third starts right from the exponential series. We will choose this path for the derivation, but we shall later see (by comparing coefficients) that we have effectively solved the integrals. The reason for not solving these directly is that one encounters a number of special cases for every configuration that appears at least twice in the path. This makes it difficult to formulate a closed expression based on the integrals.

We have (splitting H into $H = D + Y$ and sorting by the order K of Y)

$$e^{-\beta H} = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} (D+Y)^n = \sum_{K=0}^{\infty} \sum_{l=K}^{\infty} \frac{(-\beta)^l}{l!} \sum_{i_0+\cdots+i_K=l-K} D^{i_0} Y D^{i_1} \ldots D^{i_{K-1}} Y D^{i_K}.$$

Inserting identity operators yields

$$Z = \operatorname{Tr} e^{-\beta H} = \sum_{K=0}^{\infty} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} \sum_{l=K}^{\infty} \frac{(-\beta)^l}{l!} \prod_{i=1}^{K} Y_{n_{i-1},n_i} \sum_{i_0+\cdots+i_K=l-K} \prod_{j=0}^{K} D_{n_j}^{i_j}.$$

Now, have a look at the last term. We will first define the *complete homogeneous symmetric polynomial of degree l in k variables* by

$$h_l(X_1, \ldots, X_k) := \sum_{i_1+\cdots+i_k=l} \prod_{j=1}^{k} X_j^{i_j}.$$

The first few of these are given below ($X := X_1$, $Y := X_2$).

	$l=0$	$l=1$	$l=2$	$l=3$
$k=1$	1	X	X^2	X^3
$k=2$	1	$X+Y$	X^2+XY+Y^2	$X^3+X^2Y+XY^2+Y^3$

.

We therefore get

$$Z = \sum_{K=0}^{\infty} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} \prod_{i=1}^{K} Y_{n_{i-1},n_i} \sum_{l=K}^{\infty} \frac{(-\beta)^l}{l!} h_{l-K}(D_{n_0}, \ldots, D_{n_K}). \tag{5.1}$$

In general one can find $D_{n_i} = D_{n_j}$ for some $i \neq j \in \{0, \ldots, K\}$. To be precise, this always occurs at least once: We already know that, by definition, $D_{n_0} = D_{n_K} (= D_n)$. But let's first assume they are all different (we will later take the limit where some factors are equal). It is then known [25, Theorem 1.2.1] that for $l \geq K$

$$h_{l-K}(X_0, \ldots, X_K) = \sum_{i=0}^{K} \frac{X_i^l}{\prod_{j \neq i} (X_i - X_j)}. \tag{5.2}$$

As a mathematical trick and to simplify our further calculations, we will now use this expression as definition of h_{l-K} for $0 \leq l < K$ (hence, negative indices on h). The following lemma shows that the so defined terms actually vanish, which will later enable us to simplify Eq. (5.1).

Lemma 5.1. *For $0 \leq l < K$, it holds that $h_{l-K}(X_0, \ldots, X_K) \equiv 0$.*

Proof. Let $K \geq 1$ and $0 \leq l < K$. By definition,

$$h_{l-K}(X_0, \ldots, X_K) = \sum_{i=0}^{K} \frac{X_i^l}{\prod_{j \neq i} (X_i - X_j)}.$$

For $X_1, \ldots, X_K$ fixed, the first summand can be viewed as a rational function in X_0:

$$f(X_0) := \frac{X_0^l}{\prod_{j=1}^{K} (X_0 - X_j)}.$$

Since $l < K$ and all X_j, $j \in \{1, \ldots, K\}$ differ from each other, we can apply the theorem of partial fraction decomposition to obtain c_i, $i \in \{1, \ldots, K\}$ with

$$\frac{X_0^l}{\prod_{j=1}^{K} (X_0 - X_j)} = f(X_0) = \sum_{i=1}^{K} \frac{c_i}{X_0 - X_i}.$$

From the right hand side one can immediately see a way to get the coefficients, namely

$$c_i = \lim_{X_0 \to X_i} (X_0 - X_i) f(X_0) = \lim_{X_0 \to X_i} \frac{X_0^l}{\prod\limits_{\substack{j=1 \\ j \neq i}}^{K} (X_0 - X_j)} = \frac{X_i^l}{\prod\limits_{\substack{j=1 \\ j \neq i}}^{K} (X_i - X_j)}.$$

Substituting these coefficients into the partial fraction decomposition results in

$$\frac{X_0^l}{\prod\limits_{j=1}^{K} (X_0 - X_j)} = -\sum_{i=1}^{K} \frac{c_i}{X_i - X_0} = -\sum_{i=1}^{K} \frac{X_i^l}{\prod\limits_{j \neq i} (X_i - X_j)}.$$

Therefore,

$$h_{l-K}(X_0, \ldots, X_K) = \sum_{i=0}^{K} \frac{X_i^l}{\prod\limits_{j \neq i} (X_i - X_j)} = \frac{X_0^l}{\prod\limits_{j=1}^{K} (X_0 - X_j)} + \sum_{i=1}^{K} \frac{X_i^l}{\prod\limits_{j \neq i} (X_i - X_j)} = 0.$$

□

This lemma implies we can expand the following summation (which occurs in Eq. (5.1)) from $l = K$ to $l = 0$ and hence further rearrange the term.

$$\begin{aligned} \sum_{l=K}^{\infty} \frac{(-\beta)^l}{l!} h_{l-K}(X_0, \ldots, X_K) &= \sum_{l=0}^{\infty} \frac{(-\beta)^l}{l!} h_{l-K}(X_0, \ldots, X_K) \\ &= \sum_{i=0}^{K} \frac{1}{\prod_{j \neq i} (X_i - X_j)} \sum_{l=0}^{\infty} \frac{(-\beta X_i)^l}{l!} \\ &= \sum_{i=0}^{K} \frac{e^{-\beta X_i}}{\prod_{j \neq i} (X_i - X_j)}. \end{aligned} \tag{5.3}$$

If all D_{n_i} were different, we could substitute this directly into the partition sum. Unfortunately this is not the case, so that it looks like we would encounter a zero in the denominator of Eq. (5.2) for $D_{n_i} = D_{n_j}$ $(i \neq j)$. This singularity must be artificial, as one can verify by looking at the definition of the complete homogeneous symmetric polynomials. To obtain the correct value, one has to carefully take the limit $D_{n_i} \to D_{n_j}$.

Lemma 5.2. *Let $f : (X_0 - \varepsilon, X_0 + \varepsilon) \to \mathbb{R}$ be a smooth function, $l \geq 0$. Then,*

$$\lim_{X_1, \ldots, X_l \to X_0} \sum_{i=0}^{l} \frac{f(X_i)}{\prod\limits_{\substack{j=0 \\ j \neq i}}^{l} (X_i - X_j)} = \frac{1}{l!} f^{(l)}(X_0).$$

Proof. Assume $l > 0$ ($l = 0$ is clear by definition). First, Taylor-expand $f(X_i)$ around

X_0 up to order l:

$$f(X_i) = f(X_0) + f'(X_0)\,(X_i - X_0) + \cdots + \frac{1}{l!} f^{(l)}(X_0)\,(X_i - X_0)^l + \mathcal{O}\left((X_i - X_0)^{l+1}\right).$$

Collecting all those terms, we will now look at each derivative order k individually.
$\underline{k = 0}$: These terms taken together amount to

$$\sum_{i=0}^{l} \frac{f(X_0)}{\prod\limits_{\substack{j=0 \\ j\neq i}}^{l} (X_i - X_j)} = f(X_0) \sum_{i=0}^{l} \frac{1}{\prod\limits_{\substack{j=0 \\ j\neq i}}^{l} (X_i - X_j)} = f(X_0)\; h_{0-l}(X_0, \ldots, X_l) = 0$$

by Lemma 5.1.

$\underline{0 < k < l}$: The first term $i = 0$ does not contain any derivative of f, the rest stays pretty much the same.

$$\sum_{i=1}^{l} \frac{\frac{1}{k!} f^{(k)}(X_0)\; (X_i - X_0)^k}{\prod\limits_{\substack{j=0 \\ j\neq i}}^{l} (X_i - X_j)} = \frac{f^{(k)}(X_0)}{k!} \underbrace{\sum_{i=1}^{l} \frac{(X_i - X_0)^{k-1}}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{l} (X_i - X_j)}}_{=0} = 0.$$

One way to see that the last term is zero is the following: First, expand

$$(X_i - X_0)^{k-1} = \sum_{j=0}^{k-1} \binom{k-1}{j} X_i^j\, (-X_0)^{k-1-j}.$$

We then have

$$\frac{f^{(k)}(X_0)}{k!} \sum_{j=0}^{k-1} \binom{k-1}{j} (-X_0)^{k-1-j} \underbrace{\sum_{i=1}^{l} \frac{X_i^j}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{l} (X_i - X_j)}}_{=h_{j-(l-1)}(X_1,\ldots,X_l)=0} = 0,$$

again by Lemma 5.1 since $j \leq k - 1 < l - 1$.

$\underline{k = l}$: We can expand $(X_i - X_0)^{l-1}$ in the same way as before using the Binomial coefficients. We again see that all terms X_i^j where $j < l - 1$ vanish. What remains is the case $j = l - 1$, or

$$\frac{f^{(l)}(X_0)}{l!} \underbrace{\sum_{i=1}^{l} \frac{X_i^{l-1}}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{l} (X_i - X_j)}}_{=h_{l-1-(l-1)}(X_1,\ldots,X_l)=1} = \frac{f^{(l)}(X_0)}{l!},$$

since $h_0(X_1, \ldots, X_l) = 1$.

$\underline{k > l :}$ These finally are the terms that do not vanish by themselves but through the limiting procedure instead. Again, we have

$$\frac{f^{(k)}(X_0)}{k!} \sum_{i=1}^{l} \frac{(X_i - X_0)^{k-1}}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{l} (X_i - X_j)}.$$

For every single term it then follows by evaluating the limits successively that

$$\begin{aligned}
\lim_{X_1,\ldots,X_l \to X_0} \frac{(X_i - X_0)^{k-1}}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{l} (X_i - X_j)} &= \lim_{X_i,\ldots,X_l \to X_0} \frac{(X_i - X_0)^{k-1}}{(X_i - X_0)^{i-1} \prod\limits_{j=i+1}^{l} (X_i - X_j)} \\
&= \lim_{X_i,\ldots,X_l \to X_0} \frac{(X_i - X_0)^{k-i}}{\prod\limits_{j=i+1}^{l} (X_i - X_j)} \\
&= \lim_{X_{i+1},\ldots,X_l \to X_0} 0 = 0.
\end{aligned}$$

Here, it came into play that $k > l \geq i$, therefore $k - i > 0$ and the limit $X_i \to X_0$ ensures the term to become zero. □

We are now ready to look at the canonical partition function a final time. For a given path $\mathcal{C} = (K, n, n_1, \ldots, n_{K-1})$, let M be the number of distinct D_{n_i}'s and l_j, $j \in \{1, \ldots, M\}$ their multiplicity. Denote their value by D_j. This is also demonstrated in 7.3.1.

We have already shown in Eq. (5.3) that

$$\sum_{l=K}^{\infty} \frac{(-\beta)^l}{l!} h_{l-K}(X_0, \ldots, X_K) = \sum_{i=0}^{K} \frac{e^{-\beta X_i}}{\prod\limits_{\substack{j=0 \\ j\neq i}}^{K} (X_i - X_j)}. \tag{5.4}$$

Since the complete homogeneous symmetric polynomials are symmetric, we can swap the order of the X_i's as we like (this property is also found in the classical continuous time partition function). Only the offdiagonal parts Y depend on the actual order of the configuration states. First we partition $(X_0, \ldots, X_K)$ into $(X_1^1, \ldots, X_1^{l_1}, \ldots, X_M^1, \ldots, X_M^{l_M})$. For every i in the summation in Eq. (5.4) we can already insert all X_j in the denominator for which $X_i \neq X_j$: Define for $k \in \{1, \ldots, M\}$

$$f_k(X) := \frac{e^{-\beta X}}{\prod\limits_{\substack{j=1 \\ j\neq k}}^{M} (X - D_j)^{l_j}}.$$

We can then write Eq. (5.4) also as

$$\sum_{k=1}^{M}\sum_{i=1}^{l_k}\frac{f_k(X_k^i)}{\prod\limits_{\substack{j=1\\j\neq i}}^{l_k}(X_k^i-X_k^j)}.$$

Now, to evaluate at $X_k^i = D_k$ for every $k \in \{1,\dots,M\}, i \in \{1,\dots,l_k\}$, we may insert $X_k^1 = D_k$ and then investigate the limit $X_k^2,\dots,X_k^{l_k} \to D_k$. Apply Lemma 5.2 to obtain (after an index shift)

$$\sum_{l=K}^{\infty}\frac{(-\beta)^l}{l!}h_{l-K}(D_{n_0},\dots,D_{n_K}) = \sum_{k=1}^{M}\frac{1}{(l_k-1)!}f_k^{(l_k-1)}(D_k) \equiv \sum_{k=1}^{M}T[f_k,D_k,l_k-1]. \quad (5.5)$$

Here, $T[f,x_0,k]$ denotes the k-th Taylor coefficient of f around x_0 (starting at $k=0$), i.e. $T[f,x_0,k] = \frac{1}{k!}f^{(k)}(x_0)$ (see 7.3.2).

5.2 Taylor-coefficient formulation of the partition function

The last expression from Eq. (5.5) is what we can finally substitute into Eq. (5.1) to obtain

Theorem 5.3. *The canonical partition function Z can be expressed as*

$$Z = \sum_{K=0}^{\infty}\sum_{n}\sum_{n_1}\cdots\sum_{n_{K-1}}\sum_{k=1}^{M}T[f_k,D_k,l_k-1]\prod_{i=1}^{K}Y_{n_{i-1},n_i}. \quad (5.6)$$

Here, as before, $M = |\{D_{n_i} : i \in \{0,\dots,K\}\}| = |\{D_k, k \in \{1,\dots,M\}\}|$ and $l_k = |\{i \in \{0,\dots,K\} : D_{n_i} = D_k\}|$ as well as

$$f_k(X) := \frac{e^{-\beta X}}{\prod\limits_{\substack{j=1\\j\neq k}}^{M}(X-D_j)^{l_j}}. \quad (5.7)$$

With simpler configurations $\mathcal{C} = (K, n_1, \dots, n_{K-1})$ and with corresponding weights

$$W(\mathcal{C}) = \prod_{i=1}^{K}Y_{n_{i-1},n_i}\ \sum_{k=1}^{M}T[f_k,D_k,l_k-1] \quad (5.8)$$

we can write

$$Z = \sum\!\!\!\!\!\!\int_{\mathcal{C}} W(\mathcal{C}).$$

By now it should be clear that, in the title of part II, the emphasis should be laid on *Simplified Configurations* PIMC rather than *Simplified* CPIMC.

Remark 5.4. Another way to define configurations would be to split up the sum over $k \in \{1, \dots, M\}$ and express paths through $\mathcal{C}' = (K, k, n_1, \dots, n_{K-1})$ and weights

$$W(\mathcal{C}') = T[f_k, D_k, l_k - 1] \prod_{i=1}^{K} Y_{n_{i-1}, n_i}.$$

This has the benefit of simplifying later calculations, e.g. those of observables, and therefore should be taken into consideration.

Remark 5.5. If we now compare the resulting partition function with the classical continuous time CPIMC version from Chapter 3, Eq. (3.1), we see that it must also hold that

$$(-1)^K \int_0^\beta d\tau_1 \cdots \int_{\tau_{K-1}}^\beta d\tau_K \; e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1} - \tau_i)} = \sum_{k=1}^{M} T[f_k, D_k, l_k - 1]$$

(which can also be verified for some exemplary test cases). We have therefore also solved the integrals.

5.3 Reexpressing the Taylor coefficients

Up to now, in a simulation the main task would be to evaluate Taylor coefficients, which may be done numerically, though this approach would result in problems when numerically differentiating up to very high orders.

A further improvement we can make is the analytical expression of the Taylor coefficients, which circumvents the relatively low maximum order possible to achieve through numerical differentiation.

Recall from Eq. (5.7) the definition

$$f_k(X) = \frac{e^{-\beta X}}{\prod\limits_{\substack{j=1 \\ j \neq k}}^{M} (X - D_j)^{l_j}} = \frac{e^{-\beta X}}{\prod\limits_{\substack{j=0 \\ D_{n_j} \neq D_k}}^{K} (X - D_{n_j})}.$$

Another way to get $T[f_k, D_k, l_k - 1]$ is the following: By definition, if we can express f_k as

$$f_k(X) = \sum_{j=0}^{\infty} a_j (X - D_k)^j,$$

we have $T[f_k, D_k, l_k - 1] = a_{l_k-1}$. First, write

$$\begin{aligned} f_k(X) &= e^{-\beta D_k} \frac{e^{-\beta(X-D_k)}}{\prod_j (X - D_k + D_k - D_{n_j})} \\ &= e^{-\beta D_k} e^{-\beta(X-D_k)} \frac{1}{\prod_j (D_k - D_{n_j})} \prod_j \frac{1}{1 - \frac{X-D_k}{D_{n_j} - D_k}} \\ &= \frac{e^{-\beta D_k}}{\prod_j (D_k - D_{n_j})} e^{-\beta(X-D_k)} \prod_j \sum_{i_j=0}^{\infty} \left(\frac{X - D_k}{D_{n_j} - D_k} \right)^{i_j}. \end{aligned}$$

Here we have used the geometric series for values of X near to D_k. We are still looking to express f_k as a power series in $(X - D_k)$, hence we express the exponential factor as

$$e^{-\beta(X-D_k)} = \sum_{l=0}^{\infty} \frac{(-\beta)^l}{l!} (X - D_k)^l.$$

Also, the whole latter product term can be written using homogeneous polynomials as

$$\begin{aligned} \prod_j \sum_{i_j=0}^{\infty} \frac{(X - D_k)^{i_j}}{(D_{n_j} - D_k)^{i_j}} &= \sum_{i_0=0}^{\infty} \sum_{i_1=0}^{\infty} \dots (X - D_k)^{i_0+i_1+\dots} \prod_j \frac{(-1)^{i_j}}{(D_k - D_{n_j})^{i_j}} \\ &= \sum_{l=0}^{\infty} (-1)^l (X - D_k)^l \sum_{i_0+i_1+\dots=l} \prod_j \frac{1}{(D_k - D_{n_j})^{i_j}} \\ &= \sum_{l=0}^{\infty} (-1)^l (X - D_k)^l h_l \left(\frac{1}{D_k - D_{n_0}}, \frac{1}{D_k - D_{n_1}}, \dots \right) \\ &= \sum_{l=0}^{\infty} (-1)^l \, h_l(I_{k,0}, I_{k,1}, \dots) \, (X - D_k)^l. \end{aligned}$$

In the last step we have defined $I_{k,j} := \frac{1}{D_k - D_{n_j}}$. The entries of h_l are all $I_{k,j}$ which are finite, i.e. one entry for every $j \in \{0, \dots, K\}$ with $D_{n_j} \neq D_k$. In total, this means that we have $K - l_k$ possible values of j.

Last, we multiply the previous two terms and subsequently extract the coefficient of order $(X - D_k)^{l_k - 1}$:

$$T[f_k, D_k, l_k - 1] = e^{-\beta D_k} \left(\prod_{\substack{j=0 \\ D_{n_j} \neq D_k}}^{K} I_{k,j} \right) \sum_{l=0}^{l_k-1} \frac{\beta^{l_k-1-l}}{(l_k - 1 - l)!} h_l(I_{k,0}, I_{k,1}, \dots). \tag{5.9}$$

How this expression could look for a simple path is also demonstrated in 7.3.3.

Chapter 6

Observables

The formulation discussed here necessarily comes with its own estimators for observables. These are usually derived using derivatives of the partition function. The problem is that the f_k themselves can depend on this variable but also the expansion point D_k. This is the case if all the D_j are dependent on the variable in question, as in Example 6.3.

Let us now analyze this behavior in more detail. Consider a smooth function $f(x, \alpha)$. We want to take the derivative with respect to α of $T[f(\cdot, \alpha), x_0(\alpha), k]$ for some k. That is, we want to calculate

$$\frac{\partial}{\partial \alpha} \frac{1}{k!} \left(\frac{\partial}{\partial x} \right)^k \Big|_{x=x_0(\alpha)} f(x, \alpha) = \frac{1}{k!} \partial_\alpha \left[(\partial_x^k f)(x_0(\alpha), \alpha) \right].$$

This can be done by the multidimensional chain rule and yields

$$\begin{aligned} &\frac{1}{k!} \left[(\partial_x^{k+1} f)(x_0(\alpha), \alpha) \cdot x_0'(\alpha) + (\partial_\alpha \partial_x^k f)(x_0(\alpha), \alpha) \right] \\ =&\frac{1}{k!} \partial_x^k \Big|_{x=x_0(\alpha)} \left[(\partial_x f)(x, \alpha) \cdot x_0'(\alpha) + (\partial_\alpha f)(x, \alpha) \right]. \end{aligned}$$

Here we applied Schwarz's theorem to interchange the derivatives of α and x. Also, as $x_0'(\alpha)$ is seen as a constant by ∂_x, we can leave it in the domain of ∂_x^k. The result itself looks like a Taylor coefficient again: We find

$$\partial_\alpha T[f(\cdot, \alpha), x_0(\alpha), k] = T[(\partial_x f)(\cdot, \alpha) \cdot x_0'(\alpha) + (\partial_\alpha f)(\cdot, \alpha), x_0(\alpha), k]. \tag{6.1}$$

Example 6.1 Average energy. As a first example, consider the average energy. We have

$$\langle H \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{1}{Z} \sum_{\mathcal{C}} \frac{\partial W(\mathcal{C})}{\partial \beta}.$$

The only dependence of $W(\mathcal{C})$ on the temperature is in the $f_k = f_k(X, \beta)$. Since the

expansion point D_k does not depend on β itself, we have $D_k'(\beta) = 0$. Also, $\partial_\beta f_k(X,\beta) = -X f_k(X,\beta)$ by Eq. (5.7). Therefore, by Eq. (6.1),

$$\partial_\beta T[f_k(\cdot,\beta), D_k, l_k - 1] = T[0 + (\partial_\beta f_k)(\cdot,\beta), D_k, l_k - 1] = T[-X\, f_k(X,\beta), D_k, l_k - 1].$$

Inserting this into the above expression for $\langle H\rangle$ results in

$$\langle H\rangle = \frac{1}{Z}\sum\!\!\!\!\!\!\int_{\mathcal{C}} \left(\frac{\sum_{k=1}^{M} T[X\, f_k(X), D_k, l_k - 1]}{\sum_{k=1}^{M} T[f_k(X), D_k, l_k - 1]} \right) W(\mathcal{C}).$$

Choosing the paths proposed in Remark 5.4 would simplify $\langle H\rangle$ to

$$\langle H\rangle = \frac{1}{Z}\sum\!\!\!\!\!\!\int_{\mathcal{C}'} \left(\frac{T[X f_k(X), D_k, l_k - 1]}{T[f_k(X), D_k, l_k - 1]} \right) W(\mathcal{C}').$$

Even further, since $X = (X - D_k) + D_k$ and $T[(X-a)f(X), a, n] = T[f(X), a, n-1]$,

$$\langle H\rangle = \frac{1}{Z}\sum\!\!\!\!\!\!\int_{\mathcal{C}'} E(\mathcal{C}')W(\mathcal{C}')$$

with

$$E(\mathcal{C}') = D_k + \frac{T[f_k, D_k, l_k - 2]}{T[f_k, D_k, l_k - 1]}.$$

This is demonstrated in 7.3.4.

Example 6.2 Heat capacity at constant volume. The heat capacity at constant volume is related to the average energy by

$$C_V = \frac{\partial}{\partial T}\langle H\rangle\Big|_V = -\beta^2 \frac{\partial}{\partial \beta}\langle H\rangle\Big|_V$$

(setting the Boltzmann constant $k_B = 1$). Therefore,

$$\begin{aligned} C_V &= -\beta^2 \left(\frac{1}{Z^2}(\partial_\beta Z)^2 - \frac{1}{Z}\partial_\beta^2 Z \right) \\ &= -\beta^2 \langle H\rangle^2 + \frac{\beta^2}{Z}\sum\!\!\!\!\!\!\int_{\mathcal{C}'} \partial_\beta^2 W(\mathcal{C}') \\ &= \frac{1}{Z}\sum\!\!\!\!\!\!\int_{\mathcal{C}'} C(\mathcal{C}')W(\mathcal{C}') \end{aligned}$$

with

$$C(\mathcal{C}') = \beta^2 D_k^2 + \beta^2 2 D_k \frac{T[f_k, D_k, l_k - 2]}{T[f_k, D_k, l_k - 1]} + \beta^2 \frac{T[f_k, D_k, l_k - 3]}{T[f_k, D_k, l_k - 1]} - \beta^2 \langle H\rangle^2 .$$

Note that we have only derived here the expression with paths from Remark 5.4.

Example 6.3 Average occupation number. Next we will investigate the average occupation numbers. For a given configuration n_i denote the occupation number of the p-th state by $n_{i,p}$. With the expression in second quantization $H = \sum_{i,j} h_{ij} a_i^\dagger a_j + \frac{1}{2}\sum_{i,j,k,l} w_{ijkl} a_i^\dagger a_j^\dagger a_l a_k$ we have

$$\langle n_p \rangle = \frac{1}{Z}\frac{-1}{\beta}\partial_{h_{pp}} Z.$$

Only the diagonal part D of H depends on h_{pp} and is given by $\partial_{h_{pp}} D_{n_i} = n_{i,p}$.

Again, we will look at $f_k = f_k(X, h_{pp})$. This case is a little bit more complex than the previous one since f_k depends on h_{pp} (all the D_j in f_k depend on h_{pp}) and D_k itself as well. We have by Eq. (5.7)

$$(\partial_X f_k)(X, h_{pp}) = f_k(X, h_{pp}) \cdot \left[-\beta - \sum_{\substack{j=1 \\ j\neq k}}^{M} \frac{l_j}{X - D_j} \right].$$

As said earlier, $D_k'(h_{pp}) = n_{k,p}$. Last, we have

$$(\partial_{h_{pp}} f_k)(X, h_{pp}) = f_k(X, h_{pp}) \cdot \left[\sum_{\substack{j=1 \\ j\neq k}}^{M} \frac{l_j}{X - D_j} \cdot \underbrace{D_j'(h_{pp})}_{=n_{j,p}} \right].$$

Eq. (6.1) then shows that (using $\alpha := h_{pp}$)

$$\langle n_p \rangle = \frac{1}{Z} \oint_{\mathcal{C}} \left(\frac{\sum_{k=1}^{M} T[N_{k,p}(X)\, f_k(X), D_k, l_k - 1]}{\sum_{k=1}^{M} T[f_k(X), D_k, l_k - 1]} \right) W(\mathcal{C})$$

with

$$N_{k,p}(X) := n_{k,p} + \frac{1}{\beta} \sum_{\substack{j=1 \\ j\neq k}}^{M} \frac{l_j (n_{k,p} - n_{j,p})}{X - D_j}.$$

The physical interpretation of this estimator could be as follows: The average occupation number is roughly impacted by the size of the fluctuations $(n_{k,p} - n_{j,p})$, weighted by how "near" its energy D_j is to D_k. A close energy level with small fluctuations can be as important as a distant energy level with much higher fluctuations.

Chapter 7

Further Investigations

7.1 Analytical reformulation

We repeat the partition function from Theorem 5.3:

$$Z = \sum_{K=0}^{\infty} \sum_{n} \sum_{n_1} \cdots \sum_{n_{K-1}} \sum_{k=1}^{M} T[f_k, D_k, l_k - 1] \prod_{i=1}^{K} Y_{n_{i-1},n_i}. \tag{7.1}$$

This can be further modified using complex analysis. By Cauchy's integral formula one can also write

$$T[f_k, D_k, l_k - 1] = \frac{1}{(l_k-1)!} f_k^{(l_k-1)}(D_k) = \frac{1}{2\pi i} \int_{\partial B_{\varepsilon_k}(D_k)} \frac{f_k(z)}{(z-D_k)^{l_k}}\, dz$$

where $\partial B_{\varepsilon_k}(D_k)$ is the standard contour integral around D_k with some radius $\varepsilon_k > 0$, i.e. the integral along the curve $\gamma : [0, 2\pi] \to \mathbb{C}, t \mapsto D_k + \varepsilon_k\, e^{it}$. Note that

$$\frac{f_k(z)}{(z-D_k)^{l_k}} = \frac{1}{(z-D_k)^{l_k}} \frac{e^{-\beta z}}{\prod\limits_{\substack{j=1 \\ j\neq k}}^{M} (z-D_j)^{l_j}} = \frac{e^{-\beta z}}{\prod\limits_{j=1}^{M} (z-D_j)^{l_j}} = \frac{e^{-\beta z}}{\prod\limits_{l=0}^{K} (z-D_{n_l})} =: f(z)$$

where we have reversed the partition of the D_{n_i}'s in the last step. Also note that this does not depend on the k anymore. Therefore, we can write

$$T[f_k, D_k, l_k - 1] = \frac{1}{2\pi i} \int_{\partial B_{\varepsilon_k}(D_k)} f(z)\, dz =: \mathrm{Res}(f, D_k).$$

We will now assume that all D_k are bounded from below by some $-D \in \mathbb{R}$, i.e. $D_k > -D$ for all k. Consider the integration path $\Gamma_{R,-D} := [-D + iR, -D - iR] \oplus \gamma_{R,-D}$, where we first integrate from $-D + iR$ to $-D - iR$ and then along $\gamma_{R,-D} : [-\pi/2, \pi/2] \to \mathbb{C}, t \mapsto$

$-D + R\,e^{it}$ in the complex plane. By the residue theorem we have

$$\lim_{R\to\infty} \frac{1}{2\pi i} \int_{\Gamma_R} f(z)\,dz = \sum_{z\in\mathbb{C},\mathrm{Re}\,z>0} \mathrm{Res}(f,z) = \sum_{k=1}^{M} \mathrm{Res}(f, D_k) = \sum_{k=1}^{M} T[f_k, D_k, l_k - 1]$$

since the D_k are the only non-zero residues of f. Also,

$$\begin{aligned}\lim_{R\to\infty} \left| \int_{\gamma_R} f(z)\,dz \right| &\le \lim_{R\to\infty} \int_{\gamma_R} \frac{|e^{-\beta z}|}{\prod_{l=0}^{K} |z - D_{n_l}|} |dz| \\ &\le \lim_{R\to\infty} \int_{-\pi/2}^{\pi/2} \frac{e^{\beta D - \beta R\cos(t)}}{\prod_{l=0}^{K} (R - D - D_{n_l})} R\,dt \to 0\end{aligned}$$

by the dominated convergence theorem. Therefore, only the $[-D + iR, -D - iR]$ path remains:

$$\begin{aligned}\sum_{k=1}^{M} T[f_k, D_k, l_k - 1] &= \lim_{R\to\infty} \frac{1}{2\pi i} \int_{[-D+iR,-D-iR]} f(z)dz = \frac{-1}{2\pi i} \int_{-D-i\infty}^{-D+i\infty} f(z)\ \ dz \\ &= \frac{-1}{2\pi i} \int_{-D-i\infty}^{-D+i\infty} \frac{e^{-\beta z}}{\prod_{l=0}^{K} (z - D_{n_l})}\ \ dz \\ &= \frac{(-1)^{K+2}}{2\pi i} \int_{D-i\infty}^{D+i\infty} \frac{e^{\beta z}}{\prod_{l=0}^{K} (z + D_{n_l})}\ \ dz \\ &= (-1)^K \mathcal{L}^{-1}\{F(s)\}(\beta)\end{aligned}$$

with the inverse Laplace transform $\mathcal{L}^{-1}$ by the Bromwich integral formula and

$$F(s) := \frac{1}{\prod_{l=0}^{K} (s + D_{n_l})}.$$

If we compare this with the integral expression

$$\sum_{k=1}^{M} T[f_k, D_k, l_k - 1] = (-1)^K \int_0^\beta d\tau_1 \cdots \int_{\tau_{K-1}}^\beta d\tau_K e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1} - \tau_i)},$$

we find that this is given by a rather simple inverse Laplace transform. Namely,

$$\int_0^\beta d\tau_1 \cdots \int_{\tau_{K-1}}^\beta d\tau_K e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1} - \tau_i)} = \mathcal{L}^{-1}\{F(s)\}(\beta).$$

This could have also been shown, or seen, in another way: With the Heaviside step function $\Theta(t)$ one finds

$$\mathcal{L}^{-1}\left\{\frac{1}{s+D_{n_i}}\right\}(\beta) = \Theta(\beta)e^{-\beta D_{n_i}} =: e_i(\beta),$$

so that by the convolution theorem the integral expression is equal to $(e_0 * e_1 * \ldots * e_K)(\beta)$. This is easily demonstrated at the representative example of $K = 1$ kinks:

$$\begin{aligned}(e_0 * e_1)(\beta) &= \int_{-\infty}^{\infty} d\tau_1 e_0(\tau)e_1(\beta-\tau) = \int_{-\infty}^{\infty} d\tau_1 \Theta(\tau_1)e^{-D_{n_0}\tau_1}\Theta(\beta-\tau_1)e^{-D_{n_1}(\beta-\tau_1)} \\ &= \int_0^{\beta} d\tau_1 e^{-D_{n_0}\tau_1}e^{-D_{n_1}(\beta-\tau_1)} = \int_0^{\beta} d\tau_1 e^{-\sum_{i=0}^{K} D_{n_i}(\tau_{i+1}-\tau_i)}.\end{aligned}$$

That observation could have been an alternative starting point for the derivation of the partition function.

Further, this analysis may lead to a simpler Monte Carlo integration algorithm in the Laplace-transformed space, possibly evaluating also observables there. Problems lie in the difficulty of the back transformation to the real system afterwards, as this can come at a high numerical cost, as well as whether it is possible to express observables in a closed form suitable for Monte Carlo. Also, instead of one inverse temperature β, the quantities would have to be evaluated at many Laplace-transformed temperatures s in order to perform a numerical Laplace-transform inversion. On the other hand, once these were computed, one could obtain in principle the observables at all wanted system temperatures at once (compute the observables as functions of β, so to speak), though in practice this may prove to be numerically too challenging.

Another possible future investigation is the following: In the weights of each path, given by Eq. (5.8), the dependence on the length of each kink (determined through the τ_i) was eliminated. What remains is a dependence on the way the kinks are ordered, determined by the offdiagonal parts Y_{n_{i-1},n_i}. Maybe these can be grouped together so that the ordering of the kinks does not matter as well. Then, the partition function would drastically be simplified: Only different configurations and the numbers how often they occur have to be given (which in turn makes it easier to calculate the Taylor coefficients). As an example, the paths (n_0, n_1, n_2, n_0) and (n_0, n_2, n_1, n_0) can be grouped together as $\{(n_0, 2), (n_1, 1), (n_2, 1)\}$ or similarly, because their offdiagonal parts and their Taylor coefficients are equal.

7.2 Considerations for a numerical implementation

There are a number of things to consider in a possible numerical simulation. The first, and most obvious, is that in order to calculate the Taylor coefficients it seems like one has to know the multiplicities. That is, one has to know the diagonal parts D_{n_i} of each configuration n_i and has to be able to compare them in a stable way.

This could be done by choosing internal units in which these become integer or rational valued. For example, the kinetic energy can be easily expressed in an exact way if all wave vectors are $\mathbb{Z}$-valued. Since the diagonal part of the Hamiltonian usually also contains a contribution of the potential W, the diagonal part of W has to be expressable in a similar way. Alternatively, one could change the splitting $H = D + Y$ into $H = T + W$, which easens the evaluability of the Taylor coefficients as well as some Monte Carlo updates (using the update set proposed by Schoof [15]). As an example, exciting an orbital which is not part of any kink changes the weight (in the formulation of Remark 5.4) by

$$W(\mathcal{C}_2') = e^{-\beta\Delta T}\, W(\mathcal{C}_1')$$

since

$$T[f_k'(X), T_k + \Delta T, l_k - 1] = T[f_k'(X + \Delta T), T_k, l_k - 1] = e^{-\beta\Delta T}\, T[f_k, T_k, l_k - 1]$$

because all T_i change by the same amount ΔT and therefore $(X + \Delta T) - T_i' = X - T_i$. The drawback of this method would be the need to implement $T0$-Kinks, i.e. what was previously Y_{n_{i-1},n_i} can now have a diagonal part.

Alternatively, one can maybe determine whether two diagonal parts are equal solely based on the configuration, though this is unlikely with increasing degeneracy of energy modes.

Of course there are plenty of other ways for a numerical implementation. E.g., one could store the maximum diagonal part of a configuration and use fixed point numbers between 0 and 1. But the last possibility we are going to consider is particularly interesting because of its analytic background.

As shown in the previous section, the Taylor coefficients resemble the residues of a function f:

$$\sum_{k=1}^{M} T[f_k, D_k, l_k - 1] = \sum_{k=1}^{M} \mathrm{Res}(f, D_k).$$

Since f as defined before is meromorphic, one can expect a continuous dependence of f on the D_i and hence also a continuous dependence of the residues. This means, as an example, that paths $(D_{n_0}, \ldots, D_{n_i}, \ldots, D_{n_K})$ and $(D_{n_0}, \ldots, D_{n_i} + d\varepsilon, \ldots, D_{n_K})$ should

have similar weights for small $d\varepsilon$. In other words, even though there is a conceptual difference (different multiplicites and derivative orders), the numerical difference may be neglegible.

Last, we will look at what types of Hamiltonians could benefit the most from this formulation or are expected to have a certain impact on its performance. Since the Taylor coefficients mainly depend on the differences of diagonal parts, it could be an advantage to consider Hamiltonians where the diagonal parts of the potential (not the whole Hamiltonian) depend little on the given occupation, that is, are highly degenerate. Also, the degeneracy of the whole diagonal part of the Hamiltonian likely has an impact, as it is important for the number and order of the Taylor coefficients. That said, degeneracy of the energy levels of the Hamiltonian seems to be a factor which could be highly influential on the efficiency of the algorithm.

7.3 An exemplary configuration path

This section is intended as a read-along to the previous chapters. Everything mentioned here will have been referenced to earlier.

7.3.1 Notation. We will first look at the notation for a concrete path. Suppose we have a path with 3 Kinks ($K = 3$). Hence, there are three different configurations: $n = n_0 = n_3$, n_1 and n_2. Let the diagonal parts of the Hamiltonian be $D_{n_0} = D_{n_3} = a$, $D_{n_1} = b$ and $D_{n_2} = a$. Note that, although $n_2 \neq n_3$, they can still have the same diagonal part. An example for this would be momenta k and $-k$ in the ideal Fermi gas.
Continuing the investigation, we have for the number of distinct D_{n_i}'s

$$M = |\{D_{n_i} : i \in \{0, \dots, 3\}\}| = |\{a, b\}| = 2.$$

Define $D_1 := a$ and $D_2 := b$. Their corresponding multiplicities are $l_1 = 3$ and $l_2 = 1$.

7.3.2 Taylor coefficients. Next, we will calculate the Taylor coefficients. With l_i and D_i as in 7.3.1, we have

$$f_1(X) = \frac{e^{-\beta X}}{(X - D_2)^{l_2}} = \frac{e^{-\beta X}}{X - b},$$
$$f_2(X) = \frac{e^{-\beta X}}{(X - D_1)^{l_1}} = \frac{e^{-\beta X}}{(X - a)^3}.$$

Therefore,

$$T[f_1, D_1, l_1 - 1] = T[f_1, a, 2] = \frac{1}{2} f_1''(a) = \frac{e^{-\beta a}}{a - b} \left[\frac{1}{(a - b)^2} + \frac{\beta}{a - b} + \frac{\beta^2}{2} \right],$$

$$T[f_2, D_2, l_2 - 1] = T[f_2, b, 0] = f_2(b) = \frac{e^{-\beta b}}{(b-a)^3}.$$

This implies for the integrals in the partition function Eq. (3.1) (to repeat, $D_{n_0} = D_{n_2} = D_{n_3} = a$, $D_{n_1} = b$)

$$\begin{aligned}&(-1)^3 \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \int_{\tau_2}^\beta d\tau_3 e^{-D_{n_0}(\tau_1 - 0) - D_{n_1}(\tau_2 - \tau_1) - D_{n_2}(\tau_3 - \tau_2) - D_{n_3}(\beta - \tau_3)} \\ &= T[f_1, D_1, l_1 - 1] + T[f_2, D_2, l_2 - 2] = \frac{e^{-\beta a} - e^{-\beta b}}{(a-b)^3} + \beta \frac{e^{-\beta a}}{(a-b)^2} + \frac{\beta^2}{2} \frac{e^{-\beta a}}{a-b}.\end{aligned}$$

One can check that this is indeed equal to the result obtained by direct integration.

7.3.3 Reexpression of Taylor coefficients. $T[f_2, b, 0]$ can be directly calculated. The differentiation for $X = a$, as demonstrated in 7.3.2, can also be retrieved from the formula in Eq. (5.9). First, one has

$$I_{1,j} = \frac{1}{D_1 - D_{n_j}} = \frac{1}{a - D_{n_j}},$$

so

$$I_{1,0} = I_{1,1} = I_{1,3} = \frac{1}{a-a} = \infty, \quad I_{1,2} = \frac{1}{a-b}.$$

By Eq. (5.9) this means that

$$\begin{aligned}T[f_1, a, 2] &= e^{-ba}\, I_{1,2} \sum_{l=0}^{2} \frac{\beta^{2-l}}{(2-l)!} h_l(I_{1,2}) = \frac{e^{-\beta a}}{a-b} \sum_{l=0}^{2} \frac{\beta^{2-l}}{(2-l)!} I_{1,2}^l \\ &= \frac{e^{-\beta a}}{a-b} \left[\frac{\beta^2}{2} + \frac{\beta}{a-b} + \frac{1}{(a-b)^2} \right],\end{aligned}$$

which is in perfect agreement with the result from 7.3.2.

7.3.4 Observables. We now want to compute $\langle H \rangle$. We will consider paths $\mathcal{C}'$ instead of $\mathcal{C}$ as described in Remark 5.4. In Example 6.1, $E(\mathcal{C}')$ was shown to be

$$E(\mathcal{C}') = D_k + \frac{T[f_k, D_k, l_k - 2]}{T[f_k, D_k, l_k - 1]}.$$

For the path $\mathcal{C}' = (K = 3, k = 1, n_0, \ldots, n_2)$ with diagonal parts as before this implies

$$E(\mathcal{C}') = a + \frac{T[f_1, a, 1]}{T[f_1, a, 2]} = a + \frac{\beta + \frac{1}{a-b}}{\frac{\beta^2}{2} + \frac{\beta}{a-b} + \frac{1}{(a-b)^2}}.$$

The path $\mathcal{C}' = (K = 3, k = 2, n_0, \ldots, n_2)$, on the other hand, has

$$E(\mathcal{C}') = b + \frac{T[f_2, b, -1]}{T[f_2, b, 0]} = b.$$

Chapter 8

Conclusions

In this bachelor's thesis we have investigated various approaches to further improve the Configuration Path Integral Monte Carlo algorithm by Schoof et al. [14].

It was shown that the use of higher order approximations does not induce a notable change in the continuous time version of the CPIMC formulation of the canonical partition function, whereas it could make the discrete time algorithm more efficient. On the other hand, a similar approach was done for Path Integral Monte Carlo (PIMC) in coordinate representation which resulted in the Permutation Blocking PIMC algorithm [17, 18]. There, the algorithm depended explicitly on the finite length of imaginary time slices β/M, splitting them into smaller, non-equidistant slices which were evaluated differently and subsequently grouped together. The higher order approximations, in some sense, provide a CPIMC analogue to Permutation Blocking PIMC, though less permutations are grouped together. A difference is that the sub-slices are multiplied by a complex instead of a real factor and therefore cannot be physically seen as slices in imaginary time anymore. The need for this is ultimately due to the fact that it allows to circumvent the problem that the offdiagonal part Y of the Hamiltonian cannot be easily evaluated up to arbitrary order Y^n and must be split instead.

The second part was concerned with whether there is a way to reduce the configuration space in continuous time Configuration Path Integral Monte Carlo. Reducing the number of states (or kinks) in all configurations is unlikely as this seems to always involve analytically calculating sums over all possible states. On the other hand, the imaginary times $\tau_1, \ldots, \tau_K$ which occur in configurations are artificial, as in principle all possible time slice configurations are integrated over. We have therefore addressed the question whether the integration could be performed in a reasonably laborious way.

This turns out to be the case and may improve the accuracy of further simulations as well as simplifying the Monte Carlo update types, while simultaneously complicating weight and observable calculations. At the current point this modification seems unlikely

to have an impact on the fermion sign problem as the paths which are grouped together through integrating are all of the same sign.

There was also found a connection to the Laplace transform of a rather simple-looking rational function, which may provide an entry point for further analysis and perhaps even a resulting alternative algorithm. For example, using associativity and commutativity properties of the convolution operator could enable another regrouping of terms in the continuous time CPIMC canonical partition function. Alternatively, performing Monte Carlo simulations in the Laplace-transformed space potentially provides a tool to extrapolate to, or evaluate, an increased range of temperatures.

Nevertheless, the applicability and possible impact are to be investigated by future numerical simulations.

Bibliography

[1] Vladimir E. Fortov. "Extreme states of matter on Earth and in space". In: *Physics-Uspekhi* (2009). DOI: 10.3367/ufne.0179.200906h.0653.

[2] S. H. Glenzer et al. "Matter under extreme conditions experiments at the Linac Coherent Light Source". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* (2016). DOI: 10.1088/0953-4075/49/9/092001.

[3] M. D. Knudson et al. "Probing the Interiors of the Ice Giants: Shock Compression of Water to 700 GPa and 3.8 $\mathbf{g}/\mathrm{cm}^3$". In: *Phys. Rev. Lett.* (2012). DOI: 10.1103/PhysRevLett.108.091102.

[4] B. Militzer et al. "A Massive Core in Jupiter Predicted from First-Principles Simulations". In: *The Astrophysical Journal* (2008). DOI: 10.1086/594364.

[5] Pierre-François Loos and Peter M. W. Gill. "The uniform electron gas". In: *WIREs Computational Molecular Science* (2016). DOI: 10.1002/wcms.1257.

[6] Tobias Dornheim, Simon Groth, and Michael Bonitz. "The uniform electron gas at warm dense matter conditions". In: *Physics Reports* (2018). DOI: https://doi.org/10.1016/j.physrep.2018.04.001.

[7] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* (1965). DOI: 10.1103/PhysRev.140.A1133.

[8] R. O. Jones. "Density functional theory: Its origins, rise to prominence, and future". In: *Rev. Mod. Phys.* (2015). DOI: 10.1103/RevModPhys.87.897.

[9] Kieron Burke. "Perspective on density functional theory". In: *The Journal of Chemical Physics* (2012). DOI: 10.1063/1.4704546.

[10] D. M. Ceperley. "Path integrals in the theory of condensed helium". In: *Rev. Mod. Phys.* (1995). DOI: 10.1103/RevModPhys.67.279.

[11] E. Y. Loh et al. "Sign problem in the numerical simulation of many-electron systems". In: *Phys. Rev. B* (1990). DOI: 10.1103/PhysRevB.41.9301.

[12] Matthias Troyer and Uwe-Jens Wiese. "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations". In: *Phys. Rev. Lett.* (2005). DOI: 10.1103/PhysRevLett.94.170201.

[13] Tobias Dornheim et al. “Ab initio quantum Monte Carlo simulation of the warm dense electron gas”. In: *Physics of Plasmas* (2017). DOI: 10.1063/1.4977920.

[14] Tim Schoof et al. “Configuration Path Integral Monte Carlo”. In: *Contributions to Plasma Physics* (2011). DOI: 10.1002/ctpp.201100012.

[15] Tim Schoof. “Configuration Path Integral Monte Carlo: Ab initio simulations of fermions in the warm dense matter regime”. PhD thesis. Christian-Albrechts-Universität zu Kiel, 2017.

[16] Tim Schoof, Simon Groth, and Michael Bonitz. “Introduction to Configuration Path Integral Monte Carlo”. In: *Complex Plasmas: Scientific Challenges and Technological Opportunities*. Springer International Publishing, 2014, pp. 153–194. ISBN: 978-3-319-05437-7. DOI: 10.1007/978-3-319-05437-7_5.

[17] Tobias Dornheim et al. “Permutation blocking path integral Monte Carlo: a highly efficient approach to the simulation of strongly degenerate non-ideal fermions”. In: *New Journal of Physics* (2015). DOI: 10.1088/1367-2630/17/7/073017.

[18] Tobias Dornheim et al. “Permutation blocking path integral Monte Carlo approach to the uniform electron gas at finite temperature”. In: *The Journal of Chemical Physics* (2015). DOI: 10.1063/1.4936145.

[19] Franz Schwabl. *Quantenmechanik für Fortgeschrittene (QM II)*. Springer Berlin Heidelberg, 2004. DOI: 10.1007/978-3-662-09631-4.

[20] Nicholas Metropolis et al. *Equation of state calculations by fast computing machines*. Tech. rep. 1953. DOI: 10.2172/4390578.

[21] W. K. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* (1970). DOI: 10.1093/biomet/57.1.97.

[22] Tim Schoof et al. “Ab Initio Thermodynamic Results for the Degenerate Electron Gas at Finite Temperature”. In: *Phys. Rev. Lett.* (2015). DOI: 10.1103/PhysRevLett.115.130402.

[23] Siu A. Chin and C. R. Chen. “Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials”. In: *The Journal of Chemical Physics* (2002). DOI: 10.1063/1.1485725.

[24] K. Sakkos, J. Casulleras, and J. Boronat. “High order Chin actions in path integral Monte Carlo”. In: *The Journal of Chemical Physics* (2009). DOI: 10.1063/1.3143522.

[25] George M. Phillips. *Interpolation and Approximation by Polynomials*. Springer New York, 2003. DOI: 10.1007/b97417.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Weiterhin versichere ich, dass diese Arbeit noch nicht als Abschlussarbeit an anderer Stelle vorgelegen hat.

20.09.21 Paul Luis Röhrl