WIGNER FORMALISM OF QUANTUM MECHANICS AND SAMPLING OF THE WIGNER DENSITY

Antonino Emanuele Scurria

June 2023

1 Introduction

In this work, we focus on finding efficient techniques to sample from the Wigner density, an important probability density for quantum phenomena. This probability density has a very important role in computing quantum expectations of observables, recasting quantum averages as classical averages in statistical mechanics. Yet, we have to face important computational difficulties related to different mathematical areas that lead to an increasing cost. In the first part of this work, we will briefly describe the Wigner formalism of quantum mechanics. In the second part, we will focus on the computational procedure for the sampling of the Wigner density, outlining developed techniques and algorithms and proposing a new sampling algorithm that reduces the computational cost.

1.1 Wigner formalism of quantum mechanics

The Wigner formalism, named after the Hungarian physicist Eugene Wigner, is a powerful mathematical framework used in quantum mechanics to describe the behaviour of particles and systems. Developed in the early 20th century, the formalism provides a deep understanding of quantum mechanics by exploring the symmetries and transformations associated with physical systems.

At its core, the Wigner formalism introduces a unique representation of quantum states known as Wigner functions. These functions, unlike traditional wavefunctions, provide a phase-space description of quantum systems, combining elements of both position and momentum. By incorporating this duality, the Wigner formalism bridges the gap between classical and quantum mechanics, allowing us to analyse quantum phenomena within the framework of classical phase space.

One of the fundamental aspects of the Wigner formalism is its emphasis on the concept of observables. In quantum mechanics, observables are quantities that can be measured, such as position, momentum, or angular momentum. The Wigner formalism provides a mathematical framework to calculate expectation values and probabilities associated with these observables, enabling the prediction and analysis of experimental results.

Recalling that in the classical quantum approach we express the probability density in the coordinate basis as $\rho(x) = |\psi(x)|^2$ where $\psi(x) = \langle x \mid \psi \rangle$ (bra-ket notation for the scalar product) we could also think at expressing this probability amplitude in the momentum basis as $|\varphi(p)|^2$ where $\varphi(p) = \frac{1}{\sqrt{h}} \int e^{-ixp/\hbar} \psi(x) dx = \langle p \mid \psi \rangle$. Inspired then by classical

mechanics we could think about a probabilistic formalism analogous to the classical statiscal mechanics formalism, representing each quantum state as a function of both coordinates and momenta. To do so we first introduce the Weyl's transform of an operator $\hat{A}$:

$$\widetilde{A}(x,p) = \int e^{-ipy/\hbar} \langle x + y/2 | \hat{A} | x - y/2 \rangle dy \tag{1}$$

or equivalently

$$\widetilde{A}(x,p) = \int e^{-ixu/\hbar} \langle p + u/2 | \hat{A} | p - u/2 \rangle du \tag{2}$$

A key property of the Weyl's transform is that

$$\mathrm{Tr}[\hat{A}\hat{B}] = \frac{1}{h} \iint \tilde{A}(x,p)\tilde{B}(x,p) dx\, dp \tag{3}$$

Moreover if we then introduce the density operator $\hat{\rho}$ of a pure state $|\psi\rangle$ as $\hat{\rho} = |\psi\rangle\langle\psi|$ we can express the expected value of an operator as $\mathrm{Tr}[\hat{\rho}\hat{A}] = \mathrm{Tr}[|\psi\rangle\langle\psi|\hat{A}] = \langle\psi\hat{A}\psi\rangle = \langle A\rangle$. Then, exploiting 3, we get:

$$\langle \hat{A} \rangle = \mathrm{Tr}[\hat{\rho}\hat{A}] = \frac{1}{h} \int \tilde{\rho}\tilde{A} dx\, dp \tag{4}$$

Now we introduce the Wigner density of a state ψ (in the coordinate basis; the expression obtained using the Weyl's transform in the momentum basis is analogous) :

$$W(x,p) = \tilde{p}/h = \frac{1}{h} \int e^{-ipy/\hbar} \psi(x + y/2)\psi^*(x - y/2) dy \tag{5}$$

Then if we want to calculate the average of an operator $\langle \hat{A} \rangle$, exploiting the previous equations, we get:

$$\langle \hat{A} \rangle = \int \int W(x,p)\tilde{A}(x,p) dx\, dp \tag{6}$$

Now we understand the beauty and the usefulness of the Wigner formalism: exploiting the Weyl's transform and the Wigner density we can express quantum averages in a way that resembles classical statistical physics (average in the phase space). Furthermore, exploiting the fact that $\int e^{ipx/\hbar} dp = h\delta(x)$, we see that

$$\int W(x,p) dp = \psi^*(x)\psi(x).$$

$$\int W(x,p) dx = \varphi^*(p)\varphi(p)$$

Since we want to interpret the Wigner density $W(q,p)$ as a probability density we have to check the normalization condition:

$$\iint W(x,p) dx dp = \mathrm{Tr}[\hat{\rho}] = 1 \tag{7}$$

Moreover from the definition of the density operator we have that

$$\hat{\rho}^2 = \hat{\rho} \implies \mathrm{Tr}[\hat{\rho}^2] = \mathrm{Tr}[\hat{\rho}] = 1 \implies \iint W(x,p)^2 dxdp = h^{-1} \tag{8}$$

The second condition to interpret the Wigner density as a probability density to perform sampling strategies for the calculation of our observables is to ensure that $W(q,p)$ is almost everywhere positive. Unfortunately this is not the case since if we pick two different states ψ_a and ψ_b and then calculating $\mathrm{Tr}[\hat{\rho}_a\hat{\rho}_b] = |\langle\psi_a|\psi_b\rangle|^2$ we have that

$$\iint W_a(x,p)W_b(x,p)dxdp = h^{-1}\left|\langle\psi_a \mid \psi_b\rangle\right|^2$$

and considering two orthogonal states we get $\iint W_a(x,p)W_b(x,p)dxdp = 0$: Wigner densities are not almost everywhere positive. Moreover we have a bound on the modulus of $W(x,p)$: recalling that

$$W(x,p) = \frac{1}{h}\int e^{-ipy/\hbar}\psi(x+y/2)\psi^*(x-y/2)dy$$

and

$$\int \psi(x-y/2)\psi^*(x-y/2)dy = 2\int \psi(x-y/2)\psi^*(x-y/2)d(y/2) = 2$$

and picking two normalized states $\psi_1(y) = e^{-iyp/\hbar}\psi(x+y/2)/\sqrt{2}$ and $\psi_2(y) = \psi(x-y/2)/\sqrt{2}$ we get:

$$W(x,p) = 2/h\int \psi_1(y)\psi_2(y)^* dy \Rightarrow |W(x,p)| \leq 2/h$$

. Another important aspect to underline is that the Weyl's transform is invertible: this is important because it gives a on-to-one correspondence between the classical quantum formalism of wavefunctions and the Wigner formalism of Wigner densities, namely:

$$\int W(x,p)e^{ipx'/\hbar}dp = \psi\left(x+x'/2\right)\psi^*\left(x-x'/2\right)$$

and imposing $x = x'$ with a rescaling factor we get:

$$\psi(x) = \frac{1}{\psi^*(0)}\int W(x/2,p)e^{ipx/\hbar}dp$$

where $\psi^*(0)$ can be determined imposing a normalization condition. We are going now to introduce a key property of the Wigner formalism: if we consider an Hamiltonian dynamics the Hamiltonian $\hat{H}(\hat{x},\hat{p})$ we would like to calculate its average exploiting the Wigner formalism. If we consider a generic operator that is function only of the coordinate operator $\hat{A} = A(\hat{x})$ or the momentum operator $\hat{B} = B(\hat{p})$ we get :

$$\begin{aligned}\tilde{A} &= \int e^{-ipy/\hbar}\langle x+y/2|A(\hat{x})|x-y/2\rangle dy \\ &= \int e^{-ipy/\hbar}A(x-y/2)\delta(y)dy = A(x)\end{aligned} \tag{9}$$

$$\begin{aligned}\tilde{B} &= \int e^{-ixu/\hbar}\langle p+u/2|B(\hat{p})|p-u/2\rangle du \\ &= \int e^{-ixu/\hbar}B(p-u/2)\delta(u)du = B(p).\end{aligned} \tag{10}$$

Thus considering the Hamiltonian operator (naturally the weyl's transform is linear since we exploiting the linearity of the integral)

$$H(\hat{x},\hat{p}) = K(\hat{p}) + U(\hat{x}) \rightarrow \tilde{H}(x,p) = K(p) + U(x) \tag{11}$$

This implies:

$$\langle x\rangle = \iint W(x,p)x dx\, dp \qquad \langle p\rangle = \iint W(x,p)p dx\, dp$$

$$\langle K\rangle = \iint W(x,p)K(p) dx\, dp \quad \langle U\rangle = \iint W(x,p)U(x) dx\, dp$$

1.2 Harmonic oscillator in the Wigner formalism

As a practical example we consider the case of the simple quantum harmonic oscillator. In this case the Hamiltonian reads :

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2$$

. Morver the ground state (the eigenstate of the Hamiltonian operator associated to the smallest eigenvalue a.k.a. ground energy) and the first excited state (the eigenstate of the Hamiltonian operator associated to the first eigenvalue larger than the ground energy) read:

$$\psi_0(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{a}}e^{-x^2/(2a^2)}$$

$$\psi_1(x) = \frac{1}{\sqrt[4]{\pi}}\sqrt{\frac{2}{a}}\frac{x}{a}e^{-x^2/(2a^2)}$$

Thus computing the associated Wigner functions through equation 5 leads to:

$$W_0(x,p) = \frac{2}{h}\exp\left(-a^2p^2/\hbar^2 - x^2/a^2\right)$$

$$W_1(x,p) = \frac{2}{h}\left(-1 + 2(ap/\hbar)^2 + 2(x/a)^2\right)\exp\left(-a^2p^2/\hbar^2 - x^2/a^2\right)$$

Now we can compute the expected values of the energy for these first two levels using 11 :

$$\langle H\rangle = \iint W_0(x,p)\left(\frac{p^2}{2m} + \frac{m\omega^2x^2}{2}\right)dx\, dp = \frac{\hbar\omega}{2}$$

obtaining the classical result for the ground state energy of the quantum harmonic oscillator. Since in the ground state this is the only value that can be assumed it is interesting to evaluate the second momentum of the Hamiltonian operator and thus its variance:

$$\left\langle H^2\right\rangle = \iint W_0(x,p)\widetilde{\hat{H}^2}dx\, dp$$

At first we notice that $\widetilde{\hat{H}^2} = \widetilde{\hat{p}^4}/(2m)^2 + m^2\omega^4\widetilde{\hat{x}^4}/4 + \omega^2(\widetilde{\hat{x}^2\hat{p}^2} + \widetilde{\hat{p}^2\hat{x}^2})/4$. Now it is easy to calculate the first two terms; the cross products $\widetilde{\hat{x}^2\hat{p}^2} + \widetilde{\hat{p}^2\hat{x}^2}$ need instead more work :

$$\begin{aligned}\widetilde{\hat{p}^2\hat{x}^2} + \widetilde{\hat{x}^2\hat{p}^2} &= \iint p'^2 e^{-ipv/\hbar}\left(2x^2 + v^2/2\right)\left\langle x + v/2 \mid p'\right\rangle \\ &\quad \times \left\langle p' \mid x - v/2\right\rangle dv\, dp' \\ &= -\frac{1}{h}\iint e^{-ipv/\hbar}\left(2x^2 + v^2/2\right)\frac{\partial^2}{\partial v^2}e^{ip'v/\hbar}dvdp'\end{aligned}$$

Then performing a two integrations by parts on the v variable we get :

$$\begin{aligned}\widetilde{\hat{p}^2\hat{x}^2} + \widetilde{\hat{x}^2\hat{p}^2} &= -\hbar^2\int \delta(v)\frac{\partial^2}{\partial v^2}\left[e^{ipv/\hbar}\left(2x^2 + v^2/2\right)\right]dv \\ &= 2x^2p^2 - \hbar^2\end{aligned}$$

Thus we have that $\widetilde{\hat{H}^2} = H^2(x,p) - \frac{(\hbar\omega)^2}{4}$ and caarrying out the integration we obtain $\langle H^2\rangle = \frac{(\hbar\omega)^2}{4}$. Then, we finally get:

$$\Delta H = \langle H^2\rangle - \langle H\rangle^2 = \frac{(\hbar\omega)^2}{4} - \frac{(\hbar\omega)^2}{4} = 0$$

1.3 Time evolution in the Wigner formalism

In the classical quantum formalism we express the dependence on time of a state through the time dependence of its coefficient (the vectors of the basis do not depend on time). Now we want to create a correspondence between the classical quantum time evolution and the time evolution in the Wigner formalism $\psi(x,t) \longrightarrow W(x,p,t)$. Taking the derivative with respect to time of the Wigner function leads to :

$$\frac{\partial W}{\partial t} = \frac{1}{h}\int e^{-ipy/\hbar}\left[\frac{\partial\psi^*(x-y/2)}{\partial t}\psi(x+y/2) + \frac{\partial\psi(x+y/2)}{\partial t}\psi^*(x-y/2)\right]dy$$

Now, exploiting the Schrodinger equation

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar}{2im}\frac{\partial^2\psi(x,t)}{\partial x^2} + \frac{1}{i\hbar}U(x)\psi(x,t)$$

we get : $\frac{\partial W}{\partial t} = \frac{\partial W_T}{\partial t} + \frac{\partial W_U}{\partial t}$ where

$$\frac{\partial W_T}{\partial t} = \frac{1}{4\pi im}\int e^{-ipy/\hbar}\left[\frac{\partial^2\psi^*(x-y/2)}{\partial x^2}\psi(x+y/2) - \psi^*(x-y/2)\frac{\partial^2\psi(x+y/2)}{\partial x^2}\right]dy$$

$$\frac{\partial W_U}{\partial t} = \frac{2\pi}{ih^2}\int e^{-ipy/\hbar}[U(x+y/2) - U(x-y/2)]\psi^*(x-y/2)\times\psi(x+y/2)dy$$

Then, after some calculations (boring integrations by parts) we get to:

$$\frac{\partial W_T}{\partial t} = -\frac{p}{m}\frac{\partial W(x,p)}{\partial x}$$

$$\frac{\partial W_U}{\partial t} = \sum_{s=0} \left(-\hbar^2\right)^s \frac{1}{(2s+1)!} \left(\frac{1}{2}\right)^{2s} \frac{\partial^{2s+1} U(x)}{\partial x^{2s+1}} \times \left(\frac{\partial}{\partial p}\right)^{2s+1} W(x,p)$$

. The interesting aspect of this expression for the time evolution lies in the fact that the first term is purely classical while the second one takes into account non locality terms of the potential energy. Moreover we notice that, if we consider a potential energy with derivatives of order > 2 equal to 0 we get

$$\frac{\partial W_U}{\partial t} = \frac{\partial U(x)}{\partial x} \frac{\partial W(x,p)}{\partial p} \Longrightarrow$$

and

$$\frac{\partial W(x,p)}{\partial t} = -\frac{p}{m} \frac{\partial W(x,p)}{\partial x} + \frac{\partial U(x)}{\partial x} \frac{\partial W(x,p)}{\partial p}$$

that is nothing but the classical Liouville equation. As an simple example we can consider the previously introduced harmonic oscillator:

$$x_0 = x\cos(\omega t) - \frac{p}{m\omega}\sin(\omega t)$$

$$p_0 = p\cos(\omega t) + m\omega x \sin(\omega t) \Longrightarrow$$

$$W(x,p,t) = W\left(x\cos(\omega t) - \frac{p}{m\omega}\sin(\omega t), p\cos(\omega t) + m\omega x\sin(\omega t), 0\right)$$

2 Sampling strategies for the Wigner density

As in classical statistical physics for systems at thermal equilibrium the ensemble distribution is represented by the Boltzmann factor $\frac{1}{Q}e^{-\beta H}$ where β is the inverse temperature and Q is the normalization constant, also called partition function. Thus we understand that calculating the Wigner density of the Boltzmann distribution has a crucial role in the computation of ensemble averages in a quantum framework. Indeed, the application of the Wigner formalism and the Weyl transform to the Boltzann distribution allows us to write quantum expectation of observables as ensemble averages over the phase space. Moreover using 11 we understand that all the analytical and computational difficulty that we have to face in a classical quantum approach for estimating the average values of our observables is importantly reduced since we now deal with classical quantities, not operators over $L^2(\Omega)$ to be diagonalized. It is also important to underline that the path integral formalism already allows us to reduce the analytical complexity of the problem we have to face, exploiting an isomorphism with classical systems: this techniques often requires a huge computational effort. In the Wigner framework we try to recover a classical picture assuming that we are able to sample from the the thermal Wigner density. Indeed if we want to calculate the average of an observable $\langle \hat{A} \rangle$ we have:

$$\langle \hat{A} \rangle = \int dqdpW(q,p)A_w(q,p)$$

where

$$W(q,p) = \frac{1}{2\pi\hbar Q} \int \mathrm{d}\Delta e^{\frac{ip}{\hbar}\Delta} \left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle \tag{12}$$

2.1 Computational difficulties

Though at a first sight equation 12 seems to be really promising we have to take into consideration 3 important aspects :

- Lack of an analytical expression for $\left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle$
- Low numerical accuracy of integration methods due to the phase factor $e^{\frac{ip}{\hbar}\Delta}$
- adapting traditional sampling strategies with a non almost everywhere positive function (sign pproblem)
- Unknown normalization constant Q

The first problem is 'easily' solved thanks to two main tools: the first one is the so called Trotter-Suzuki decomposition; the second one is the path integral formalism developed by physicist Richard Feynman. The second and the third problem will be addressed exploiting an asymptotic probabilistic expansion called Edgeworth Expansion. The last computational problem will be faced through sampling techniques that don't require the normalization constant such as Markov Chain Monte Carlo or Langevin Dynamics that do not rquire an analytical expression for the density (i.e. Noisy Monte Carlo methods etc...).

2.2 Trotter-Suzuki decomposition and Path Integral formalism

To find a solution for the first problem of the previous subsection we have to find a way of obtaining an usable analytical expression of $\left\langle q - \frac{\Delta}{2} \left| e^{-\beta \hat{H}} \right| q + \frac{\Delta}{2} \right\rangle$: this term is nothing but a matrix element of the operator $e^{-\beta \hat{H}}$ expressed in the coordinates basis. At first, we now introduce the Trotter-Suzuki decomposition. Let us consider an Hamiltonian operator of the form $\hat{H} = \sum_{i=0}^{k} \hat{h}_i$ where $\left[\hat{h}_i, \hat{h}_j\right] \neq 0$ if $i \neq j$ ($[\cdot, \cdot]$ is the simple commutator defined as $[a, b] = ab - ba$) then we can write the Trotter-Suzuki decomposition of the first order as

$$\exp\left(-i\hat{H}\Delta_t\right) = e^{-i\hat{h}_1\Delta_t} \dots e^{-i\hat{h}_K\Delta_t} + \mathcal{O}\left(\Delta_t^2\right)$$

whereas, the second order version of the formula reads:

$$\exp\left(-i\hat{H}\Delta_t\right) = e^{-i\hat{h}_1\frac{\Delta_t}{2}} \dots e^{-i\hat{h}_K\frac{\Delta_t}{2}} e^{-i\hat{h}_K\frac{\Delta_t}{2}} \dots e^{-i\hat{h}_1\frac{\Delta_t}{2}} + \mathcal{O}\left(\Delta_t^3\right)$$

where Δ_t is small enough. Now we are going to use this fundamental tool to develop the path integral formalism: in our case we have $-i\Delta_t = -\beta$ (also called imaginary time) and at first, since we need a small imaginary time step, we are going to trivially rewrite

$$e^{-\beta\hat{H}} = \underbrace{e^{-\Delta_\tau\hat{H}} \dots e^{-\Delta_\tau\hat{H}}}_{P \text{ times}}$$

where $P \times \Delta_t = \beta$ (we can do that since $\hat{H}$ naturally commutes with itself). Then if we have two states $\mathbf{x}$ and $\mathbf{y}$ and we want to calculate the matrix element of our quantum Boltzmann factor, inserting P resolutions of the identity $\sum_{\mathbf{x}_j} |\mathbf{x}_j\rangle \langle \mathbf{x}_j| = \hat{I}$ (we are dealing with an orthonormal basis), we can write:

$$\begin{aligned}
\langle \mathbf{x} \mid e^{-\beta\hat{H}} \mid \mathbf{y} \rangle &= \langle \mathbf{x}| e^{-\Delta_\tau\hat{H}} \dots e^{-\Delta_\tau\hat{H}} |\mathbf{y}\rangle \\
&= \sum_{\mathbf{x}_0} \langle \mathbf{x}| \underbrace{e^{-\Delta_\tau\hat{H}} \dots e^{-\Delta_\tau\hat{H}}}_{P \text{ times}} |\mathbf{x}_0\rangle \langle \mathbf{x}_0 \mid \mathbf{y}\rangle \\
&= \sum_{\mathbf{x}_0} \sum_{\mathbf{x}_1} \langle \mathbf{x}| \underbrace{e^{-\Delta_\tau\hat{H}} \dots e^{-\Delta_\tau\hat{H}}}_{P-1 \text{ times}} |\mathbf{x}_1\rangle \langle \mathbf{x}_1| e^{-\Delta_\tau\hat{H}} |\mathbf{x}_0\rangle \langle \mathbf{x}_0 \mid \mathbf{y}\rangle \\
&= \sum_{\mathbf{x}_0\mathbf{x}_1\dots\mathbf{x}_{P-1}} \langle \mathbf{x}| e^{-\Delta_\tau\hat{H}} |\mathbf{x}_{P-1}\rangle \dots \langle \mathbf{x}_1| e^{-\Delta_\tau\hat{H}} |\mathbf{x}_0\rangle \langle \mathbf{x}_0 \mid \mathbf{y}\rangle .
\end{aligned}$$

Now we define the propagator as

$$G^{\Delta_\tau}(\mathbf{x}, \mathbf{y}) \equiv \left\langle \mathbf{x} \left| e^{-\Delta_\tau\hat{H}} \right| \mathbf{y} \right\rangle$$

and we can rewrite the previous equation as :

$$\langle \mathbf{x} \mid e^{-\beta\hat{H}} \mid \mathbf{y}\rangle = \sum_{\mathbf{x}_0\mathbf{x}_1\dots\mathbf{x}_{P-1}} \underbrace{G^{\Delta_\tau}(\mathbf{x}, \mathbf{x}_{P-1}) \dots G^{\Delta_\tau}(\mathbf{x}_1, \mathbf{x}_0)}_{P \text{ times}} \langle \mathbf{x}_0 \mid \mathbf{y}\rangle$$

Now we are able to use the Trotter-Suzuki decomposition (Δ_t is as small as we want)of the first and second order obtaining respectively :

$$\begin{aligned} G^{\Delta_\tau}(\mathbf{x},\mathbf{y}) &= \left\langle \mathbf{x} \left| e^{-\Delta_\tau \hat{H}_0} e^{-\Delta_\tau \hat{H}_1} \right| \mathbf{y} \right\rangle + \mathcal{O}\left(\Delta_\tau^2\right) \\ &= e^{-\Delta_\tau H_0(\mathbf{x})} \left\langle \mathbf{x} \left| e^{-\Delta_\tau \hat{H}_1} \right| \mathbf{y} \right\rangle + \mathcal{O}\left(\Delta_\tau^2\right) \\ &= e^{-\Delta_\tau H_0(\mathbf{x})} G_1^{\Delta_\tau}(\mathbf{x},\mathbf{y}) + \mathcal{O}\left(\Delta_\tau^2\right), \end{aligned}$$

$$G^{\Delta_\tau}(\mathbf{x},\mathbf{y}) = e^{-\Delta_\tau \frac{H_0(\mathbf{x})}{2}} G_1^{\Delta_\tau}(\mathbf{x},\mathbf{y}) e^{-\Delta_\tau \frac{H_0(\mathbf{y})}{2}} + \mathcal{O}\left(\Delta_\tau^3\right)$$

where we have assumed that we can split the Hamiltonian operator in two non commuting with the first one $\hat{H}_0$ that is diagonal in the coordinate basis.

In our case we are interested in an Hamiltonian operator of the form:

$$\hat{H} = -\frac{\hbar^2}{2m} \sum_i^N \nabla_{\vec{r}_i}^2 + \sum_i V_1\left(\vec{r}_i\right)$$

where N is the number of particles and we are in a 3-dimensional space. Then $G_1^{\Delta_t}$, also called 'free propagator', can be easily computed exploiting $\langle x \mid p\rangle = \langle p \mid x\rangle^* = \frac{1}{\sqrt{2\pi\hbar}} e^{ixp/\hbar}$ where $\mid x\rangle$ is an eigenvector of the coordinate basis and $\mid p\rangle$ is an eigenvector of the momentum basis. In the specific case in 12 we deal with only one particle in a one-dimensional space and we get:

$$\begin{aligned} G_1^{\Delta_\tau}(\mathbf{x},\mathbf{y}) &= \langle \mathbf{x} \mid e^{-\Delta_t \frac{\hat{p}^2}{2m}} \mid \mathbf{y}\rangle \\ &= \int dp \, \langle \mathbf{x} \mid e^{-\Delta_t \frac{\hat{p}^2}{2m}} \mid p\rangle \langle p \mid \mathbf{y}\rangle \\ &= \int dp \, e^{-\Delta_t \frac{p^2}{2m}} \frac{1}{2\pi\hbar} e^{ip(\mathbf{x}-\mathbf{y})/\hbar} \\ &= \sqrt{\frac{m}{2\pi\hbar^2\Delta_t}} e^{-\frac{m(\mathbf{x}-\mathbf{y})^2}{2\Delta_t\hbar^2}} \end{aligned}$$

In the end, applying all these passages to $\left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle$, introducing the auxiliary variables $x_1 \dots x_{\nu-1}$ we get :

$$\begin{aligned} &\left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle \propto \exp\left\{ -\frac{mv}{2\beta\hbar^2} \left[\left(q - \frac{\Delta}{2} - x_{v-1} \right)^2 \right.\right. \\ &\left.\left. + \left(q + \frac{\Delta}{2} - x_1 \right)^2 + \sum_{\lambda=1}^{v-2} \left(x_{\lambda+1} - x_\lambda \right)^2 \right] \right\} \\ &\times \exp\left\{ -\frac{\beta}{v} \left[\frac{1}{2} V\left(q + \frac{\Delta}{2} \right) + \frac{1}{2} V\left(q - \frac{\Delta}{2} \right) + \sum_{\lambda=1}^{v-1} V\left(x_\lambda \right) \right] \right\} \end{aligned}$$

Now we have an approximated analytical expression that we can use!

2.3 Graham-Charlier and Edgeworth Series

In the field of probability theory and statistics, the study of approximations to probability distributions has been of great importance. Two widely used methods for approximating probability distributions are the Graham Charlier series and the Edgeworth series. These series expansions provide valuable tools for understanding and analyzing complex distributions in terms of simpler and more manageable forms.

The Graham Charlier series, named after J. R. Graham and A. W. F. Charlier, is a powerful technique for approximating a probability density function (PDF) using a finite number of terms in a polynomial expansion. This series is particularly useful for distributions that exhibit skewness and kurtosis, allowing for a more accurate representation of the tails of the distribution compared to traditional Gaussian approximations. By incorporating higher-order moments of the distribution, the Graham Charlier series can capture the shape and characteristics of non-normal distributions with greater precision. A drawback of this approach is that we can't control the error we are neglecting and it often diverges very quickly.

On the other hand, the Edgeworth series, named after F. Y. Edgeworth, providing a method for approximating the PDF of a random variable by using a series expansion in terms of cumulants, is a true asymptotic expansion of our proabability density. Similar to the Graham Charlier series, the Edgeworth series is effective in capturing the shape and tail behavior of a distribution beyond what is possible with a Gaussian approximation. By incorporating higher-order cumulants, the Edgeworth series provides a more detailed and accurate representation of the distribution, especially in the presence of skewness.

In this subsection, we will explore the foundations and applications of the Graham Charlier and Edgeworth series. We will discuss their mathematical formulations, and then we will focus on the application on our specific problem.

2.3.1 Graham-Charlier Series

At first we are going to recall some standard definitions in probability theory. Let $F(x)$ be the cumulative probability distribution of a random variable X. Then the mean value for the random variable $g(X)$ is the expectation

$$\mathrm{E}g(X) \equiv \langle g(X) \rangle \equiv \int_{-\infty}^{\infty} g(x) \mathrm{d}F(x)$$

The probability density function, under some mild assumptions, is $p(x) = \mathrm{d}F(x)/\mathrm{d}x$. Following the definition of the mean value, the k-th order moment of X is

$$\alpha_k \equiv \mathrm{E}X^k = \int_{-\infty}^{\infty} x^k \, \mathrm{d}F(x)$$

We denote the cumulative normal distribution by

$$P(x) \equiv \int_{-\infty}^{x} Z(t) \mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-t^2/2\right) \mathrm{d}t$$

so its probability density function is the Gaussian function

$$Z(x) = \frac{\exp\left(-x^2/2\right)}{\sqrt{2\pi}}$$

We also recall the definition of orthogonal polynomials. Two polynomials $P_n(x)$ and $P_m(x)$ of degrees $n \neq m$ are orthogonal on the real axis with respect to a weight function $w(x)$ if

$$\int_{-\infty}^{\infty} w(x)P_n(x)P_m(x) = 0.$$

In our specific case we are intrested in Chebyshev-Hermite polynomials that we denote as $H_n(x)$ where the wight function is $w(x) = \exp\left(-x^2/2\right) \propto Z(x)$. According to Rodrigues' formula,

$$He_n(x) = (-1)^n \mathrm{e}^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2/2}.$$

These polynomials are very useful since when we have a probability density function $p(x)$ which is nearly Gaussian, it seems natural to use the expansion

$$p(x) \sim \sum_{n=0}^{\infty} c_n \frac{\mathrm{d}^n Z(x)}{\mathrm{d}x^n}$$

where the coefficients c_n measure the deviations of $p(x)$ from $Z(x)$. From the previous definitions it is easy to show that $\frac{\mathrm{d}^n Z(x)}{\mathrm{d}x^n} = (-1)^n He_n(x)Z(x)$, so that $p(x) \sim \sum_{n=0}^{\infty}(-1)^n c_n He_n(x)Z(x)$, with

$$c_n = \frac{(-1)^n}{n!} \int_{-\infty}^{\infty} p(t)He_n(t)\mathrm{d}t.$$

This is the well-known Gram-Charlier series (of type A). Since $He_n(x)$ is a polynomial, we understand that the coefficients c_n are a linear combination of the moments α_k of the random variable X with PDF $p(x)$. The linear combination has an analytical formula that can be derived exploiting some properties of Chebyshev-Hermite polynomials.

2.3.2 Edgeworth Series

If we have a random variable X that can be normalized then we can use the Edgeworth expansion for the normalized variable $\tilde{X}$ that unlike the Graham-Charlier Series, is a true asymptotic expansion where the error can be controlled. We now introduce the definition of the characteristic variable of a random function X:

$$\Phi(t) = \int_{-\infty}^{\infty} e^{itx} d\mathcal{F}(x)$$

that is the Fourier transform of $p(x)$ if the probability density $p(x) = \mathrm{d}F(x)/\mathrm{d}x$ exists. From the definition of the characteristic function implies that if the moment $\alpha_k(2)$ of X exists, $\Phi^{(k)}(0) = i^k\alpha_k$ This implies that we can write a Taylor series of $\Phi(t)$ as

$$\Phi(t) \sim 1 + \sum_{k=1}^{\infty} \frac{\alpha_k}{k!}(it)^k$$

. The same procedure applies alto to $\ln \Phi(t)$,

$$\ln \Phi(t) \sim \sum_{n=1}^{\infty} \frac{\kappa_n}{n!}(it)^n$$

, where

$$\kappa_n \equiv \frac{1}{i^n}\left[\frac{\mathrm{d}^n}{\mathrm{d}t^n}\ln\Phi(t)\right]_{t=0}$$

are the cumulants of our random variable.

Exploiting now a fundamental lemma for the chain derivative of $f\circ g(x)\equiv f(g(x))$, which reads

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}f(g(x)) = \\ n!\sum_{\{k_m\}} f^{(r)}(y)\Bigg|_{y=g(x)} \prod_{m=1}^{n}\frac{1}{k_m!}\left(\frac{1}{m!}g^{(m)}(x)\right)^{k_m}$$

where $r=k_1+k_2+\ldots+k_n$ and the set $\{k_m\}$ consists of all non-negative integer solutions of the Diophantine equation

$$k_1+2k_2+\ldots+nk_n=n.$$

Using this lemma a useful formula for the cumulants κ_n and the moments α_k can be derived :

$$\kappa_n = n!\sum_{\{k_m\}}(-1)^{r-1}(r-1)!\prod_{m=1}^{n}\frac{1}{k_m!}\left(\frac{\alpha_m}{m!}\right)^{k_m}.$$

Here summation extends over all non-negative integers $\{k_m\}$ satisfying the previous Diophantine equation and $r=k_1+k_2+\ldots+k_n$.

Now we consider a random variable X with $\mathrm{E}X=0$ (this can always be imposed and it is also the case for our specific problem), and let X have variance equal to σ^2. If X has the characteristic function $\Phi(t)$, then the normalized random variable X/σ has the characteristic function $\varphi(t)=\Phi(t/\sigma)$. Therefore we have that

$$\ln\varphi(t)=\ln\Phi(t/\sigma)\sim\sum_{n=2}^{\infty}\frac{\kappa_n}{\sigma^n n!}(it)^n$$

(the sum starts from index one because of the constraint on the mean value of our random variable X)

$$\varphi(t)\sim \mathrm{e}^{-t^2/2}\exp\left\{\sum_{n=3}^{\infty}\frac{S_n\sigma^{n-2}}{n!}(it)^n\right\},$$

with

$$S_n\equiv\kappa_n/\sigma^{2n-2}$$

Now we write the exponential function in as a formal series in powers of σ,

$$\exp\left\{\sum_{r=1}^{\infty}\frac{S_{r+2}\sigma^r}{(r+2)!}(it)^{r+2}\right\}\sim 1+\sum_{s=1}^{\infty}\mathcal{P}_s(it)\sigma^s,$$

where the coefficient of the power s is a function $\mathcal{P}_s(it)$. Now, using

$$g(x)\equiv\sum_{r=1}^{\infty}\left\{S_{r+2}(it)^{r+2}x^r/(r+2)!\right\}$$

and $f \equiv \exp$, and exploiting the lemma on chain derivation, we get:

$$\begin{aligned}\mathcal{P}_s(it) &\equiv \frac{1}{s!}\frac{\mathrm{d}^s}{\mathrm{d}x^s} f(g(x))\Big|_{x=0} \\ &= \sum_{\{k_m\}} \prod_{m=1}^{s} \frac{1}{k_m!}\left(\frac{S_{m+2}(it)^{m+2}}{(m+2)!}\right)^{k_m}\end{aligned}$$

where the summation extends again over all non-negative integers k_m satisfying the previous Diophantine equation.

Now that we achieved this we want to go back to the original probability density of X: the PDF for X/σ is $q(x) \equiv \sigma p(\sigma x)$, and it is the inverse Fourier transform of the characteristic function φ :

$$q(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-itx}\varphi(t)\mathrm{d}t$$

Exploiting now a useful property of the the Fourier transform of a function $p(x)$ we know that $(-it)^n\Phi(t)$ is the transform of the n-th derivative of $p(x)$,

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-itx}(-it)^n\Phi(t)\mathrm{d}t$$

Recalling that fhe Fourier transform of a gaussian distribution $Z(x)$ is $\exp\left(-t^2/2\right)$, we notice that each $(it)^n$, multiplied by $\exp\left(-t^2/2\right)$ in the expansion of φ , generates the n-th derivative of $Z(x)$,

$$(-1)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}Z(x) = \int_{-\infty}^{\infty} \mathrm{e}^{-itx}(it)^n \exp\left(-t^2/2\right)\mathrm{d}t,$$

Now we can write the expansion for $q(x)$ as:

$$\begin{aligned}q(x) &= Z(x) + \sum_{s=1}^{\infty}\sigma^s \\ &\times\left\{\sum_{\{k_m\}}\prod_{m=1}^{s}\frac{1}{k_m!}\left(\frac{S_{m+2}(-1)^{m+2}}{(m+2)!}\frac{\mathrm{d}^{m+2}}{\mathrm{d}x^{m+2}}\right)^{k_m} Z(x)\right\}.\end{aligned}$$

Here the set $\{k_m\}$ in the sum consists of all non-negative integer solutions of the equation

$$k_1 + 2k_2 + \ldots + sk_s = s.$$

Using the fact that $He_n(x) = (-1)^n\mathrm{e}^{x^2/2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-x^2/2}$. and $r = k_1+k_2+\ldots+k_s$ we can rewrite the previous formula exploiting the Chebyshev-Hermite polynomials:

$$\begin{aligned}q(x) = \sigma p(\sigma x) &= Z(x)\left\{1 + \sum_{s=1}^{\infty}\sigma^s\right. \\ &\left.\times\sum_{\{k_m\}} H_{s+2r}(x)\prod_{m=1}^{s}\frac{1}{k_m!}\left(\frac{S_{m+2}}{(m+2)!}\right)^{k_m}\right\}.\end{aligned}$$

This is the Edgeworth expansion for arbitrary order s.

2.3.3 Application of the Edgeworth expansion to the Wigner density

After having shown the theory related to Graham-Charlier Series and the Edgeworth Series it is now time to apply it to solve the second and the third problems: low numerical precision due to the phase factor $e^{\frac{ip}{\hbar}\Delta}$ and the non almost everywhere positiveness of the Wigner function that is a crucial problem for common sampling algorithms.
Let us begin by rewriting 12 as

$$W(q,p) = \frac{1}{2\pi\hbar Q} e^{-\beta U(q)} e^{-\kappa_2(q)\frac{p^2}{2\hbar^2}} \int \mathrm{d}\Delta e^{\frac{ip}{\hbar}\Delta + \kappa_2(q)\frac{p^2}{2\hbar^2}} \rho_c(\Delta \mid q),$$

where

$$e^{-\beta U(q)} = \int \mathrm{d}\Delta \left| q - \frac{\Delta}{2} \right| e^{-\beta\hat{H}} \left| q + \frac{\Delta}{2} \right\rangle,$$

$$\rho_c(\Delta \mid q) = \frac{\left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle}{\int \mathrm{d}\Delta \left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle},$$

and $\kappa_2(q)$ is, at the moment, an arbitrary function of position. In this rewriting $e^{-\beta U(q)}$ is nothing but the normalization factor of this new distribution function $\rho_c(\Delta \mid q)$. The careful reader might have noticed why we are doing so: our aim is that of interpreting the Weyl transform of our quantum Boltzmann factor as a characteristic function to then exploit the Edgeworth series we have introduced in the previous subsection. Indeed, choosing κ_2 as the second cumulant (in this case, by definition, ρ_c is an even function $\implies \langle\Delta\rangle_{\rho_c} = 0$),

$$\kappa_2(q) = \int \mathrm{d}\Delta\Delta^2 \rho_c(\Delta \mid q) = \left\langle \Delta^2 \right\rangle_{\rho_c}$$

we obtain the Edgeworth expansion for the characteristic function,

$$\begin{aligned} W(q,p) =& \frac{1}{2\pi\hbar Q} e^{-\beta U(q)} e^{-\kappa_2(q)\frac{p^2}{2\hbar^2}} \\ & \times \underbrace{\left[1 + \frac{\kappa_4(q)}{4!} \left(\frac{ip}{\hbar} \right)^4 + \frac{\kappa_6(q)}{6!} \left(\frac{ip}{\hbar} \right)^6 + \cdots \right]}_{\mathcal{C}_{EW}(q,p)} \end{aligned} \tag{13}$$

From now on we will indicate with

$$EW_0 = \frac{1}{2\pi\hbar Q} e^{-\beta U(q)} e^{-\kappa_2(q)\frac{p^2}{2\hbar^2}} \tag{14}$$

, and with EW_n the approximation that takes into account up to the n-th cumulant in the Edgeworth expansion 13. It is important to notice that waht we have done now allows us to overcome the previous two difficulties. Thanks to this expansion we don't have to deal anymore with the phase factor that is responsible for poor numerical precision. Moreover the equation 13 also helps us in the third critical problem: thanks to this expansion we can now perform a standard sampling strategy without any sign problem to calculate the cumulants we need in order to approximate our probability density function. It is also interesting to notice that if we have a Gaussian distribution the approximation EW_0 is exact. Moreover it can be seen experimentally that the approximation EW_0 is often enough to approximate $W(q,p)$.

2.4 Sampling Methods

In the previous sections we have built a framework to have analytical access to the matrix elements and we have also created a way of eliminating the phase phactor and the sign problem due to the fact that the Wigner density can assume negative values. Finally, in this subsection, we will discuss about how to sample the Wigner density in order to be able to calculate averages of observables we are interested in; indeed we recall that our final goal is a way to statistically compute avergaes of the kind:

$$\langle \hat{A} \rangle = \int dqdpW(q,p)A_w(q,p)$$

We could use different approaches to do that: one is to perform a Langevin dynamics to sample the probability distribution. In this work we are going instead towards a more classical choice that is a classical Markov Chain Monte Carlo method. The non trivial problem in our case is that we can't perform a single Monte Carlo sampling since, as shown in 13, our EW_n approximation is computed using linear combinations of moments associated to the random variable Δ with respect to the density we have introduced as $\rho_c(\Delta, q)$. The problem is indeed that for every step of the main Monte Carlo sampling of EW_n we have to perform an auxiliary Monte Carlo to calculate the cumulants we need and estimate $e^{-\beta U(q)}$. Fortunately, as previously said, the only EW_0 is often enough to capture the quantum behaviour of our systems but even in this case we have to perform an auxiliary Monte Carlo to approximate two quantities. Moreover another additional difficulty is that when performing the main Monte Carlo step we have to take into account that we are not dealing with an analytical expression for the probability density we want to estimate, but with an estimate of it. For this reason we have now to exploit a Monte Carlo algorithm that takes into account noisy estimates of our probability distribution when performing the main Monte Carlo sampling.

2.4.1 Noisy Monte Carlo method

As previously said, we need an algorithm that takes into account the fact that we don't have an analytical expression for the density but only a statistical estimate. Let us suppose that when a move from x to x' is made, we estimate the 'energy difference', which we denote $\delta(x \to x')$. By $H(x)$ we mean the true Hamiltonian (this can be generalize but we are assuming to deal with a physical system at thermal equilibrium so that our probability density is represented by the Boltzmann density). Let $a\left(x \to x'\right)$ the modified acceptance probability with the assumption that it depends only on the estimate δ of the energy difference. Let $P(\delta; x \to x')\, d\delta$ be the probability for obtaining a value δ. Then we can calculate the average acceptance from x to x' as,

$$A\left(x \to x'\right) = \int_{-\infty}^{\infty} d\delta P\left(\delta; x \to x'\right) a(\delta). \tag{15}$$

The detailed balance equation is

$$\begin{aligned} & e^{-H(x)/k_B T} T\left(x \to x'\right) A\left(x \to x'\right) \\ & \quad = e^{-H(x')/k_B T} T\left(x' \to x\right) A\left(x' \to x\right). \end{aligned} \tag{16}$$

$$\Delta\left(x \to x'\right) = \left[H\left(x'\right) - H(x)\right]/k_B T - \ln\left[T\left(x' \to x\right)/T\left(x \to x'\right)\right]$$

we can rewrite the detailed balance equation as

$$A\left(x \to x'\right) = e^{-\Delta} A\left(x' \to x\right).$$

Now we assume that δ is symmetric in x and x' then $P\left(\delta; x' \to x\right) = P\left(-\delta; x \to x'\right)$. Then detailed balance requires

$$\int_{-\infty}^{\infty} d\delta P\left(\delta; x \to x'\right)\left[a(\delta) - e^{-\Delta} a(-\delta)\right] = 0. \tag{17}$$

As one can see, the difficulty now is that during Markov Chain Monte Carlo algorithm, we do not know either $P\left(\delta; x \to x'\right)$ or Δ. Hence we must find a function $a(\delta)$ which satisfies the detailed balance condition for all $P(\delta)$ and Δ.

To overcome the difficulty of dealing with such a formula we suppose that $P\left(\delta; x \to x'\right)$: we assume that the noise of the energy difference will be normally distributed (this is reasonable because in many physical scenarios the Hamiltonian is given by a sum of single contributions and thus we can exploit the central limit theorem). Given that $\langle\delta\rangle = \Delta$, the probability of getting a particular value of δ is

$$P(\delta) = \left(2\sigma^2\pi\right)^{-1/2} \exp\left(-(\delta - \Delta)^2/\left(2\sigma^2\right)\right). \tag{18}$$

Moreover we will assume that we know the value of σ^2.

In the case of a normal distribution with known variance σ^2 the detail balance condition is given by

$$a_P(\delta; \sigma) = \min\left(1, \exp\left(-\delta - \sigma^2/2\right)\right). \tag{19}$$

From the previous formula we can thus notice that in the noisy Monte Carlo framework the acceptance probability is modified by a multiplicative factor equal to $\exp\left(-\sigma^2/2\right)$ for $\delta > -\sigma^2/2$. From this we can see that larger the noise is, less probable is to accept the trial configuration.

To prove that this acceptance formula satisfies the detailed balance condition, one does the integrals in 15 to obtain

$$A(\Delta) = \frac{1}{2}\left[e^{-\Delta} \operatorname{erfc}\left(c\left(\sigma^2/2 - \Delta\right)\right) + \operatorname{erfc}\left(c\left(\sigma^2/2 + \Delta\right)\right)\right]$$

where $\operatorname{erfc}(z)$ is the complimentary error function and $c = 1/\sqrt{2\sigma^2}$
So, to use the formalism of our problem what we need to do during the sampling procedure for the Wigner density is to estimate the value k_2 exploiting an ausiliary Markov Chain Monte Carlo over the density $\rho_c(\Delta \mid q)$; then we have access to an estimate of the final density we want to estimate $e^{-k_2 \frac{p^2}{2\hbar^2}}$. So considering a Metropolis-Hastings procedure we estimate the variance of our estimator of k_2 and apply the previous procedure.

2.5 Harmonic oscillator Sampling

As a simple example we now analyze the harmonic oscillator case. It is an interesting case because, if we consider an harmonic potential, we have an analytical expression for the matrix element $\left\langle q - \frac{\Delta}{2} \left| e^{-\beta \hat{H}} \right| q + \frac{\Delta}{2} \right\rangle$. To simplify the notation we first define the matrix element $f(q, \Delta) \equiv \left\langle q - \frac{\Delta}{2} \left| e^{-\beta \hat{H}} \right| q + \frac{\Delta}{2} \right\rangle$. As previously said we have an analytical expression for

$$f(q, \Delta) \propto \exp\left\{ -\frac{m\omega}{4\hbar} \tanh^{-1}(\beta\hbar\omega/2)\Delta^2 \right\} \exp\left\{ -\frac{m\omega}{\hbar} \tanh(\beta\hbar\omega/2) q^2 \right\}$$

so that

$$\kappa_2(q) = \frac{\int \mathrm{d}\Delta \Delta^2 \exp\left\{ -\frac{m\omega}{4\hbar} \tanh^{-1}(\beta\hbar\omega/2)\Delta^2 \right\}}{\int \mathrm{d}\Delta \exp\left\{ -\frac{m\omega}{4\hbar} \tanh^{-1}(\beta\hbar\omega/2)\Delta^2 \right\}} = \frac{2\hbar}{m\omega} \tanh(\beta\hbar\omega/2)$$

As we can see κ_2 in this case is independent of q and it is thus a constant .

Finally, note that in the harmonic oscillator case, since $f(q, \Delta)$ is Gaussian in Δ, all cumulants of order higher than two are zero, $C_{EW}(q, p) = 1$ and the approximation EW_0 is thus analytically exact. Finally the Wigner density $W(q, p)$ for the harmonic oscillator case reads:

$$W(q, p) = \frac{1}{2\pi\hbar Q} \sqrt{2\pi \frac{2\hbar}{m\omega} \tanh(\beta\hbar\omega/2)} e^{-\frac{2\hbar}{m\omega} \tanh(\beta\hbar\omega/2) \frac{p^2}{2\hbar^2}}$$

In this case the sampling strategy is much easier since we don't have to estimate k_2 and for this reason we don't have to perform a nested sampling taking into account uncertainties of our estimates.

2.6 EW_0 approximation

In this section we are going to focus more on the EW_0 approximation of our Wigner density $W(q, p)$. It is interesting to notice that the EW_0 approximation is nothing but a Laplace approximation of our initial probability density $\rho_c(\Delta \mid q)$. When we truncate the coefficient $C_{EW}(q)$ to the first order, taking into account only the cumulant of order 2 (variance) then we are left with 14: the expression of EW_0 is telling us that $\rho_c(\Delta \mid q)$ is nothing but a Gaussian of mean $\mu = 0$ and variance $\sigma^2 = k_2$ because the factor $e^{-k_2 \frac{p^2}{2\hbar^2}}$ can be interpreted as the characteristic function of $\rho_c(\Delta \mid q)$: since the characteristic function of a probability density uniquely specifies it and in our case it is the characteristic function of a Gaussian, we have that $\rho_c(\Delta \mid q)$ is a Gaussian as previously affirmed. This is interesting because we now understand that the estimate of k_2 is nothing but the calculation of the variance of a Gaussian distribution. This simplifies the estimation since the integral that we need to calculate

$$\kappa_2(q) = \int \mathrm{d}\Delta \Delta^2 \rho_c(\Delta \mid q) = \left\langle \Delta^2 \right\rangle_{\rho_c}$$

can now be calculated knowing the partition function of our induced distribution $\rho_c(\Delta \mid q)$. Indeed we have that:

$$\begin{aligned}\kappa_2(q) &= \int \mathrm{d}\Delta\Delta^2 \rho_c(\Delta \mid q) \\ &= \int \mathrm{d}\Delta\Delta^2 \frac{\left\langle q - \frac{\Delta}{2} \right| e^{-\beta\hat{H}} \left| q + \frac{\Delta}{2} \right\rangle}{\sqrt{2\pi k_2}} \\ k_2^{3/2} &= \int \mathrm{d}\Delta\Delta^2 \frac{\left\langle q - \frac{\Delta}{2} \right| e^{-\beta\hat{H}} \left| q + \frac{\Delta}{2} \right\rangle}{\sqrt{2\pi}}\end{aligned} \tag{20}$$

This implies that to calculate k_2 we don't need algorithms that allows us to neglect the normalization constant since the normalization constant itself is a function of k_2. This opens the door to exploit also different integration techniques than standard Markov Chain Monte Carlo with Metropolis-Hastings scheme: one possible and tested solution could be the one of exploiting the Bayesian Monte Carlo technique that has a faster rate of convergence compared to a standard Monte Carlo sampling. This technique was tested with unfortunately poor results due to the fact that when we want to exploit a probabilistic technique we have also to perform importance sampling with respect to our numerator i.e.

$$\frac{1}{\sqrt{2\pi}} \int \mathrm{d}\Delta\Delta^2 \frac{\left\langle q - \frac{\Delta}{2} \right| e^{-\beta\hat{H}} \left| q + \frac{\Delta}{2} \right\rangle}{p(\Delta)} p(\Delta)$$

where $p(\Delta)$is a suitable probability density (in the Bayesian Monte Carlo method the $p(\Delta)$ is chosen as a N-dimensional (N is the number of auxiliary variables + the real variable Δ) Gaussian density). Another idea would also be to use a Quasi Monte Carlo method but, even in this case we have to face the problem of finding a good function for importance sampling. Another interesting fact of the EW_0 approximation is that in this case the factor $e^{-\beta U(q)}$ can be (almost) analytically known: indeed we have just seen that $\rho_c(\Delta \mid q)$ is nothing but a Gaussian in this approximation and this implies that

$$e^{-\beta U(q)} = \sqrt{2\pi k_2}$$

Now we understand that we can sample EW_0 only estimating k_2. The careful reader might have noticed a trouble in this procedure. Let us suppose that this is true and we want to calculate the marginal probability density of $W(q,p) = W(q)$, then we have

$$\begin{aligned}W(q) &= \int_{-\infty}^{\infty} \mathrm{d}p W(q,p) \\ &= \frac{1}{2\pi\hbar Q} \int_{-\infty}^{\infty} \mathrm{d}p \sqrt{2\pi k_2} e^{-k2\frac{p^2}{2\hbar^2}} \\ &= \frac{1}{Q}\end{aligned}$$

This is naturally absurd since we get a uniform distribution over a set of infinite measure. Moreover this reasoning is general and would be valid regardless of the Wigner density we are dealing with. Indeed it is true that $\rho_c(\Delta \mid q)$ is a Gaussian distribution in this approximation but we have to notice also that

$$\left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle = C(q) \times e^{\frac{-\Delta^2}{2k_2}}$$

and

$$e^{-\beta U(q)} = \sqrt{2\pi k_2} \times C(q).$$

So we cannot directly sample the EW_0 approximation after estimating k_2 but we need also an estimate of $C(q)$.
What is now interesting to notice is that, instead of sampling the Wigner density as previously described, we could sample $W(q,p) = W_q(q) \times W_{p|q}(p)$ where $W_q(q)$ is the marginal density of the coordinates and $W_{p|q}(p)$ is the conditional probability density for the momenta. Indeed, in doing so, we can avoid the computation of the unknown term $C(q)$ since, as previously seen, integrating the Wigner density over the momenta gives

$$W_q(q) = \frac{C(q)}{Q}$$

. Then, if we are able to compute the marginal density for the coordinates we obtain

$$W_{p|q}(p) = \frac{1}{2\pi\hbar}\sqrt{2\pi k_2} e^{-k2\frac{p^2}{2\hbar^2}}.$$

Thus we would have an analytical expression for the momenta density. Now we analyze the term $W_q(q)$ and we get:

$$W_q(q) = \frac{1}{2\pi\hbar Q} \int \mathrm{d}\Delta \int \mathrm{d}p \, e^{\frac{ip}{\hbar}\Delta} \left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle \tag{21}$$

$$= \frac{1}{2\pi Q} \int \mathrm{d}\Delta \, \delta(\Delta) \left\langle q - \frac{\Delta}{2} \left| e^{-\beta\hat{H}} \right| q + \frac{\Delta}{2} \right\rangle \tag{22}$$

$$= \frac{1}{2\pi Q} \left\langle q \left| e^{-\beta\hat{H}} \right| q \right\rangle \tag{23}$$

Now we exploit the previous path integral formalism to express the diagonal elements of the operator $e^{-\beta\hat{H}}$. Once we obtain $W_q(q)$ we can then compute the marginal density for the momenta as :

$$W_p(p) = \mathbf{E}\left[W_{p|q}(p)\right]_{W_q(q)} \tag{24}$$

Naturally if we are in a low-dimensional case we can compute the density exploiting simple numerical integration techniques. Even though this procedure allows us to compute even the entire density for the momenta, we recall that the reason why we are doing all this work is to compute averages for observables we are interested in, namely:

$$\langle \hat{A} \rangle = \int \int W(x,p)\tilde{A}(x,p)dx \, dp$$

Following the last reasoning we then write

$$\langle \hat{A} \rangle = \int \int W_q(q)W_{p|q}(p)\tilde{A}(x,p)dx \, dp$$

We thus need only to sample from these two probability densities to calculate averages. The algorithm is therefor straightforward:

Algorithm 1 Sampling of the Wigner density

$k_2 = [\ ]$
$average = [\ ]$
while not converged **do**
  $q_{trial} \leftarrow new\_configuration$
  compute $R = \frac{W_q(q_{trial})}{W_q(q_i)}$
  generate $\xi \sim U[0,1]$
  if $R > \xi$ **then**
    $q_{i+1} \leftarrow q_{trial}$
  else
    $q_{i+1} \leftarrow q_i$
  end if
end while
while not converged **do**
  sample $q \sim W_q(q)$
  while not converged **do**
    $\Delta_{trial} \leftarrow new\_configuration$
    compute $R = \frac{\rho_c(\Delta_{trial}|q)}{\rho_c(\Delta_i|q)}$
    generate $\xi \sim U[0,1]$
    if $R > \xi$ **then**
      $\Delta_{i+1} \leftarrow \Delta_{trial}$
    else
      $\Delta_{i+1} \leftarrow \Delta_i$
    end if
  end while
  $k_2 = 0$
  while not converged **do**
    sample $\Delta \sim \rho_c(\Delta|q)$
    k_2.append(Δ^2)
  end while
  $k_2 = mean(k_2)$
  sample $p \sim \mathcal{N}(0, \frac{\hbar^2}{k_2})$
  average.append($A(q,p)$)
  $k_2 = 0$
end while
average = mean(average)

It is easy to notice that this algorithm massively impacts the computational complexity since in the sampling algorithms for the Wigner density that exploit nested sampling (both Langevin dynamics and Monte Carlo sampling methods) we have to compute k_2 during the thermalization process; here instead we will compute k_2 only when needed and for much less steps. Moreover we assumed to consider only symmetric transition probabilities in proposing new states. Another important aspect of the algorithm is that we are estimating the analytical expression of $W_{p|q}(p)$ through k_2: if we now call the estimate we get through the Monte Carlo method as $\tilde{k_2}$ and we assume that:

$$k_2 = \tilde{k_2} + \delta$$

where δ is a noisy term. Then to have a more accurate estimate of our $W_{p|q}(p)$ we can do the following

$$W_{p|q}(p) = \frac{1}{\sqrt{2\pi\hbar^2}} \sqrt{\tilde{k_2} + \delta} \times e^{-(\tilde{k_2}+\delta)\frac{p^2}{2\hbar^2}} \tag{25}$$

$$= \frac{1}{\sqrt{2\pi\hbar^2}} \sqrt{\tilde{k_2}} \times e^{-k_2\frac{p^2}{2\hbar^2}} \times f(q,p,\delta) \tag{26}$$

where $f(q,p,\delta)$ is equal to

$$f(q,p,\delta) = \sqrt{1 + \frac{\delta}{\tilde{k_2}}} e^{-\delta\frac{p^2}{2\hbar^2}}$$

Then, estimating the noise of our estimator $\tilde{k_2}$ through $\langle\Delta^4\rangle$ we can follow the previous algorithm just augmenting our observable $A(q,p)$ with the noisy term:

$$\tilde{A}(q,p) = A(q,p) f(q,p,\delta)$$

It is important to notice the the algorithm can be extended to high-dimensional systems taking into account that the previous k_2 will be a matrix (covariance matrix), k_4 will be a forth-order tensor and so on, namely:

$$[\kappa_2(\mathbf{q})]_{ij} = \int \mathrm{d}\Delta \Delta_i \Delta_j \rho_c(\Delta \mid \mathbf{q})$$

and

$$\begin{aligned} W(\mathbf{q},\mathbf{p}) =& \frac{1}{(2\pi\hbar)^N Q} \exp\left\{ -\frac{\mathbf{p}^\top \kappa_2(q)\mathbf{p}}{2\hbar^2} - \beta U(\mathbf{q}) \right\} \\ & \times \left(1 + \sum_{i,j,k,l} \frac{p_i p_j p_k p_l}{24\hbar^4} \Big[[\kappa_4(\mathbf{q})]_{ijkl} - [\kappa_2(\mathbf{q})]_{ij} [\kappa_2(\mathbf{q}) \right. \\ & \left. - [\kappa_2(\mathbf{q})]_{ik} [\kappa_2(\mathbf{q})]_{jl} - [\kappa_2(\mathbf{q})]_{il} [\kappa_2(\mathbf{q})]_{jk} \Big] + \cdots \right) \end{aligned}$$

with k_4 defined by

$$[\kappa_4(\mathbf{q})]_{ijkl} = \int \mathrm{d}\Delta \Delta_i \Delta_j \Delta_k \Delta_l \rho_c(\Delta \mid \mathbf{q}).$$

2.7 Experimental Results

The algorithm described above has been tested with the trivial case of the harmonic oscillator potential and the more significant Morse potential , a physically important potential for atomistic modeling. The results obtained have been compared with the Wigner density computed by discretizing the Hamiltonian operator to obtain the boltzmann density and to compute numerically the integral of the Wigner density. For the algorithm presented we focused on tuning parameters for the thermalization steps in order to fastly obtain a converged Markov Chain, and to perform statistical analysis, calculating the autocorrelation time and the statistical variance. The algorithms has been tested with different temperatures (the lower the temperature the stronger the quantum effects) obtaining constantly a good performance. For the initial sampling of the marginal coordinate density we have performed the easy Monte Carlo move of moving only a single bead for each step. For the Monte Carlo sampling to estimate k_2 we have considered the classical move of sampling a trial configuration that is uniformely distributed in an interval around the previous configuration, optimizing the length of the interval to obtain faster convergence and an acceptance rate $\sim 50\%$

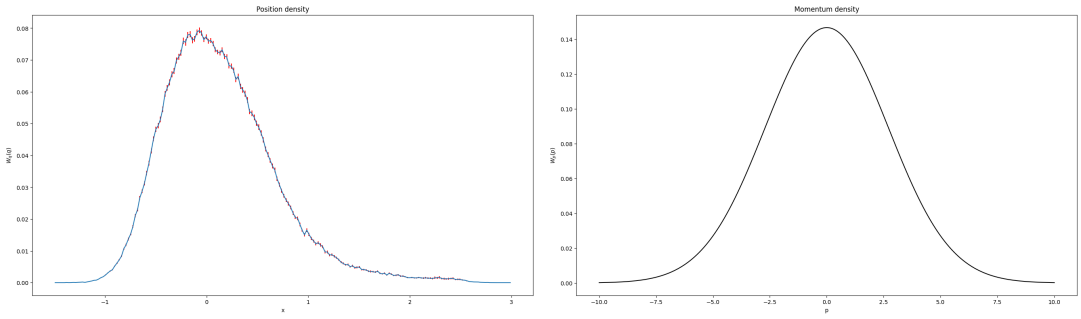

Figure 1: Plot of the coordinate density and the momentum density for the Morse Potential at 1200K obtained by the Wigner sampling algorithm. For the position density we have also computed statistical errors.

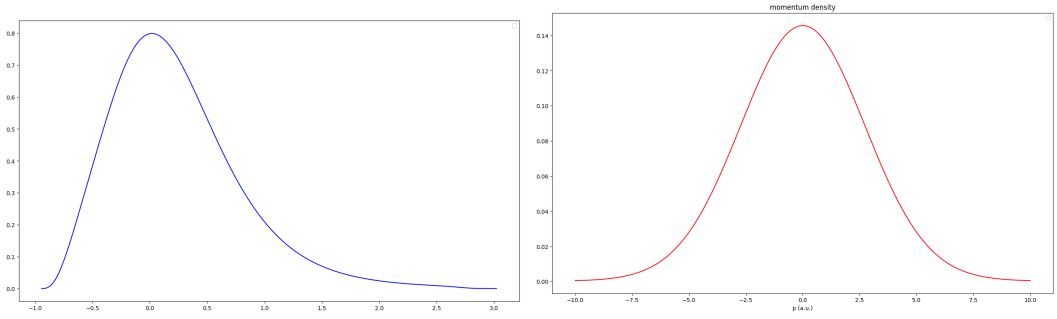

Figure 2: Plot of the coordinate density and the momentum density for the Morse Potential at 1200K obtained by discretization of the Hamiltonian operator.