Thermodynamics of phonons with the HMC — v2

Project for Computational Physics II, WS19/20   *March 6, 2020*

1 Grading scheme

Weight	Assessed competence
20%	Physical Transcription + Planning
35%	Implementation
15%	Testing
30%	Running + Numerical Analysis + Visualization + Physical Analysis

Deadline: 30.03.2020

2 Physical system

We want to study the thermodynamics of the phonons in a one-dimensional cristal. In principle one can study higher dimensional systems with the same techniques, but one would need a parallel code (look into the CPIII course!). "Phonons" is a fancy word that stands for the vibrational modes of the atoms of the cristal around their equilibrium positions.

These vibrational modes are intrisically quantum mechanical. As all quantum mechanical system, one can write a Hamiltonian operator H. Statistical mechanics teaches us that the thermodinamics of the system can be derived e.g. by the partition function of the canonical ensemble, i.e.

$$Z(\beta) = \operatorname{tr} e^{-\beta H} \,, \tag{1}$$

where $\beta = (k_\mathrm{B} T)^{-1}$ is related to the temperature T, and the trace is taken over the Hilbert space. For realistic systems it is impossible to calculate the partition function by means of this fomula, and one needs to develop smart algorithms to do the job.

We will consider an algorithm based on a different representation of the partition function, i.e. on the so-called *path-integral* formulation. In order to understand why the path-integral formulation is equivalent to the standard quantum-mechanical formulation, you will need to take specific courses. Here I will only tell you how it works.

We consider a one-dimensional cristal with N atoms, which we label with an index $n = 1, \ldots, N$. At equilibrium, the atoms are equispaced with a distance of 1 (we neglect physical units here). The atoms can vibrate for simplicity only in the longitudinal direction. We call h_n the displacement of the n-th atom from its equilibrium position.

The path-integral formalism requires the introduction of an axiliary continuous coordinate s, which can take values from 0 to β. All degrees of freedom become functions of s, in particular the displacements h_n are replaced by functions $h_n(s)$. The theory of the path-integral says that, in order to describe the canonical partition function correctly, one needs to choose *periodic boundary conditions* in the coordinate s.

The path-integral representation of the partition function is obtained from the Euclidean action of the system:

$$S(h) = \int_0^\beta ds \left\{ \frac{1}{2} \sum_{n=1}^N [\partial_t h_n(s)]^2 + \frac{v_0^2}{2} [h_1(s)]^2 + \frac{v_0^2}{2} \sum_{n=2}^N [h_n(s) - h_{n-1}(s)]^2 \right\} \tag{2}$$

Notice that the $n = 1$ term $[h_1(s)]^2$ is treated differently. If you think about it, this is equivalent to introduce another atom with $n = 0$ and requiring that $h_0 = 0$, i.e. the first atom is kept fixed. This allows to remove translational modes of the whole cristal which do not contribute to the thermodynamics. The parameter v_0 is the speed of sound (group velocity of waves at low frequency).

Notice that the index n is naturally discrete, however the extra coordinate s is continuous. As usual we need to discretize it. This can be done by dividing the interval $[0, \beta]$ in P equal intervals with length β/P, by replacing the the coordinate s with the index $p = 0, 1, 2, \ldots, P - 1$ and

$$h_n(s) \to h_{n,p} \ , \tag{3}$$

by replacing the integral in s by a sum

$$\int_0^\beta ds \to \frac{\beta}{P} \sum_{p=0}^{P-1} \ , \tag{4}$$

and by replacing the derivative in s with the discrete definition

$$\partial_t h_n(s) \to \frac{P}{\beta}(h_{n,p+1} - h_{n,p}) \ . \tag{5}$$

The discrete derivative is defined with periodic boundary conditions in p.

The theory of the path-integral says that the partition function can be calculated as

$$Z(\beta) = \lim_{P \to \infty} \int [dh] \, e^{-S(h)} \ , \tag{6}$$

where we use the short-hand notation for the following multidimensional integral

$$\int [dh] = \left[\int_{-\infty}^\infty \prod_{n=1}^N \prod_{p=0}^{P-1} dh_{n,p} \right] \ . \tag{7}$$

As for the Ising model, we will not be able to calculate the partition function. We will be able to calculate thermal expectation values of observables. For instance, the mean square fluctuation of the atom in n can be calculated as

$$\langle [h_{n,p}]^2 \rangle_\beta = \lim_{P \to \infty} \frac{1}{Z}(\beta) \int [dh] \, e^{-S(h)} [h_{n,p}]^2 \ . \tag{8}$$

The average fluctuation can be defined simply as the square root of this quantity. Apparently, the expectation value depends also on p, i.e. on the extra coordinate s. In fact one can prove that this is not the case: the action enjoys translational invariance in the extra dimension, and this implies that the expectation value $\langle [h_{n,p}]^2 \rangle_\beta$ does not depend on p. In fact one can use a better estimator for the mean square fluctuation by averaging over all values of p.

The average potential energy can be calculated as the following expectation value

$$V_\beta = \frac{v_0^2}{2} \left\langle \left\{ [h_{1,p}]^2 + \sum_{n=2}^N [h_{n,p} - h_{n-1,p}]^2 \right\} \right\rangle_\beta \ . \tag{9}$$

Again, this expactation value does not depend on p. A direct calculation of the kinetic energy K is more involved. However, in the case of quadratic action, one can use the virial theorem and conclude that $V = K$. The internal energy of the system is simply $U = K + V = 2V$.

3 Hybrid Monte Carlo (HMC)

A pure Metropolis (as for the Ising model) would be very inefficient for calculating the desired expectation values. The Hybrid Monte Carlo (HMC) is an important sampling algorithm with consists of three main ingredients:

- a thermal gaussian bath;
- a Molecular-Dynamics (MD) evolution;
- a Metropolis step.

As for the Ising model, the HMC generates a Markov chain. The statistical analysis needed to obtain expectation values with their error is the same as for the Ising model project.

The HMC algorithm needs the introduction of auxiliary degrees of freedom, the *conjugate momenta* $\pi_{n,p}$, and a fictitious Hamiltonian function

$$H_{\mathrm{MD}}(\pi, h) = \frac{1}{2} \sum_{n,p} \pi_{n,p}^2 + S(h) \ , \tag{10}$$

which is usually referred to as MD Hamiltonian. As for the Ising model, one needs to start with a given initial condition for $h_{n,p}$, for instance one can simply choose $h = 0$ or draw h randomly.

A single HMC step starts from a given configuration h and generates the next configuration in the Markov chain. The next configuration may be equal to the previous one. This is how the HMC step works.

1. **Heat-bath for momenta.** Generate each $\pi_{n,p}$ for $n = 1, \ldots, N$ and $p = 0, \ldots, P$ randomly with a gaussian distribution centered in 0 and with width equal to 1.

2. **Molecular Dynamics (MD) evolution.** Introduce an extra coordinate (yes, another one!) t, which is usually referred to as MD time, and solve the MD equations[1]

$$\frac{\partial \tilde{\pi}_{n,p}}{\partial t}(t) = F(\tilde{h}(t)) \ , \tag{11}$$

$$\frac{\partial \tilde{h}_{n,p}}{\partial t}(t) = \tilde{\pi}_{n,p}(t) \ , \tag{12}$$

   with initial condition

$$\tilde{h}_{n,p} = h_{n,p} \ , \qquad \tilde{\pi}_{n,p} = \pi_{n,p} \ , \tag{13}$$

   up to $t = 1$, and with the definition

$$F_{n,p}(h) = -\frac{\partial S}{\partial h_{n,p}}(h) \ . \tag{14}$$

   This function is usually referred to as *MD force*. Obviously the MD equations need to be solved with an approximated method. We will comment on this later.

3. **Metropolis step.** Notice that $\tilde{h}(0)$ is the old configuration. The proposed new configuration for the Markov chain is $\tilde{h}(1)$. The proposed new configuration is accepted or rejected according to a Metropolis step. Calculate the variation of the MD Hamiltonian

$$\Delta H = H(\tilde{\pi}(1), \tilde{h}(1)) - H(\tilde{\pi}(0), \tilde{h}(0)) \ . \tag{15}$$

   If $\Delta H < 0$, then accept $\tilde{h}(1)$ as the new configuration for the Markov chain. If $\Delta H > 0$, then accept $\tilde{h}(1)$ as the new configuration for the Markov chain with probability $e^{-\Delta H}$. If $\tilde{h}(1)$ is rejected, then $\tilde{h}(0)$ is repeated in the Markov chain.

---

[1] Notice that these are nothing but the Hamilton equation associated to the MD Hamiltonian.

In order to turn the above list of steps in a real algorithm, we need to specify a numerical scheme to integrate the MD equations. As usual, this is done by dividing the interval $[0, 1]$ in M equal smaller intervals with length $\tau = 1/M$ (this is the *MD time step*). We will use the simple *leapfrog integrator*, which amounts to replacing step 2 with:

2a. Set the initial conditions $\tilde{h} = h$ and $\tilde{\pi} = \pi$.

2b. Repeat the following set of instructions (in the given order) M times:

$$\tilde{\pi} \leftarrow \tilde{\pi} + \frac{\tau}{2} F(\tilde{h}) \,, \tag{16}$$

$$\tilde{h} \leftarrow \tilde{h} + \tau \tilde{\pi} \,, \tag{17}$$

$$\tilde{\pi} \leftarrow \tilde{\pi} + \frac{\tau}{2} F(\tilde{h}) \,. \tag{18}$$

In particular notice that, in the third step $F(\tilde{h})$ is calculated with the new $\tilde{h}$.

2c. At the end of the MD evolution, we identify $\tilde{h}(1) = \tilde{h}$ and $\tilde{\pi}(1) = \tilde{\pi}$.

3.1 Monitoring the HMC: acceptance rate

It is usually a good idea to keep track of the number of configurations that have been accepted along the Markov chain. The ratio of accepted configuration over total number of configurations is called *acceptance rate*. If the number M of MD time steps is very large, or equivalently if the MD time step τ is very small, the acceptance rate is very close to 100%. Obivously a single HMC step becomes computationally very expensive in this limit. On the other hand, if M is small the single HMC step becomes very cheap but the algorithm accepts only rarely, which makes the algorithm very inefficient. You will need to tune (i.e. choose after trying) a value for M for which the calculation is not too expensive, and the algorithm is decently efficient. Usually an acceptance rate in the 80%–95% range is considered a good one.

4 Project goals

- Write a very short document with the explicit expression of the discretized action $S(h)$ and of the force $F(h)$. Let me check these as soon as you have them.

- Write a code for the generation of configurations with the HMC algorithm. This code does not need to be written for a general number of dimensions (consider that the algorithm is already fairly complex, you don't want to add an extra layer of complexity). Keep the code modular. The code should simulate only one Markov chain at the time, for a given choice of parameters. Make sure that the parameters needed to run the simulations are not hard-coded and can be easily chosen at run-time (e.g. by means of command line arguments).

- The code should calculate the potential energy and the mean squared displacement of the $n = N$ atom in the cristal for each configuration and print them. The code should also print whether each Metropolis step has accepter or rejected the new configuration.

- Develop some tests for this code. I will propose one, which is compulsory, in the next point.

- (*Test of the MD evolution.*) Fix $N = 10$, $v_0 = 0.2$, $\beta = 0.7$, $P = 50$. Choose $h = 0$ and some random π as initial condition. Choose several values for $M = 10, 20, 30, \dots$, and run the MD evolution (only this, not the whole HMC step!) for each of them, always with the *same* initial condition. Calculate ΔH as a function of $\tau = 1/M$, and show that ΔH vanishes in the $\tau \to 0$ limit. If τ is small enough, ΔH is proportional to some integer power of τ. Which power?

- Generate a Markov chain for the parameters $N = 10$, $v_0 = 0.2$, $\beta = 0.7$, $P = 50$. You will need to choose M (or τ) in such a way that the acceptance rate is in the 80%–95% range. Plot the history of the observables, estimate the thermalization, calculate their average and error (you will need the autocorrelation time). Document the process with which you chose M, and also document how you have done the statistical analysis.

- Recall that one needs to take the $P \to \infty$ limit in order to recover the correct thermal averages. Repeat the simulation and analysis for $P = 100$. What can you say about the $P \to \infty$ limit of the considered observables?