Second problem (PH-353)

Monte Carlo in Quantum Mechanics

Biagio Lucini

February 2008

Pseudorandom number generators

Monte Carlo integration

Monte Carlo in Statistical

Monte Carlo in Quantum Mechanics

Analysis of Monte Carlo data

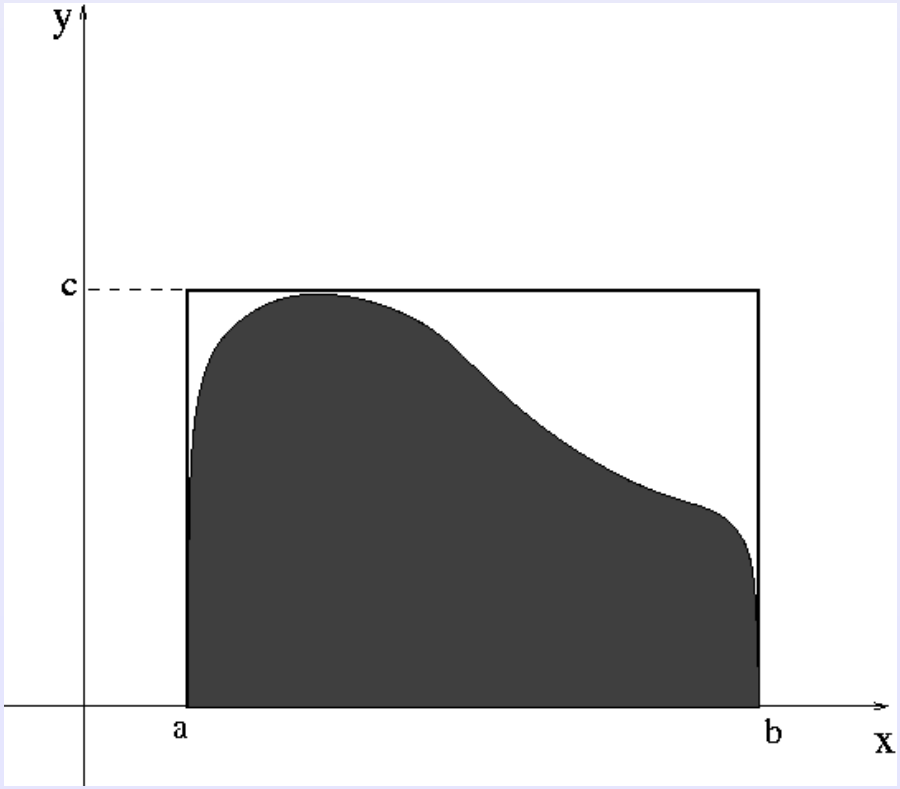
Assignment

Part I

Pseudorandom number generators

Integrals and random numbers

Simple problem: compute the integral below:



Random numbers

Physical approach: construction of a cosmic ray detector

$$\text{area} \propto \text{number of events per second}$$

We need to know the number of events per unit area $\Rightarrow$ the detector has known area

Numerical computation $\Rightarrow$ code for the generation of random numbers
Deterministic algorithm for the generation of *random* numbers

Random numbers

Physical approach: construction of a cosmic ray detector

$$\text{area} \propto \text{number of events per second}$$

We need to know the number of events per unit area $\Rightarrow$ the detector has known area

Numerical computation $\Rightarrow$ code for the generation of random numbers
Deterministic algorithm for the generation of *random* numbers

Contradiction?

Random numbers

Physical approach: construction of a cosmic ray detector

$$\text{area} \propto \text{number of events per second}$$

We need to know the number of events per unit area $\Rightarrow$ the detector has known area

Numerical computation $\Rightarrow$ code for the generation of random numbers
Deterministic algorithm for the generation of *random* numbers

Contradiction?
No: PSEUDOrandom numbers

Random number generators

Algorithm that generates a flat distribution $P(x)$ in the interval [0,1]

$$\int_a^b P(x)\mathrm{d}x = b - a, \qquad 0 \leq a \leq b < 1$$

$\bar{P}$ is then homogeneous in the interval [A,B[

$$\bar{P}(x) = A + (B - A) \times P(x)$$

Other probability distributions can be obtained from the homogeneous distribution

e.g. Box-Müller transformation $\Rightarrow$ gaussian distribution

Properties of a good generator

- Correct statistical distribution (characteristic time τ_S)
- Long period (characteristic time τ_P)
- Absence of correlations (characteristic time τ_C)
- Independent sequences for (*semi-*)independent inputs

Properties of a good generator

- Correct statistical distribution (characteristic time τ_S)
- Long period (characteristic time τ_P)
- Absence of correlations (characteristic time τ_C)
- Independent sequences for (*semi-*)independent inputs

Properties of a good generator

- Correct statistical distribution (characteristic time τ_S)
- Long period (characteristic time τ_P)
- Absence of correlations (characteristic time τ_C)
- Independent sequences for (*semi*-)independent inputs

Properties of a good generator

- Correct statistical distribution (characteristic time τ_S)
- Long period (characteristic time τ_P)
- Absence of correlations (characteristic time τ_C)
- Independent sequences for (*semi*-)independent inputs

Properties of a good generator

- Correct statistical distribution (characteristic time τ_S)
- Long period (characteristic time τ_P)
- Absence of correlations (characteristic time τ_C)
- Independent sequences for (*semi*-)independent inputs

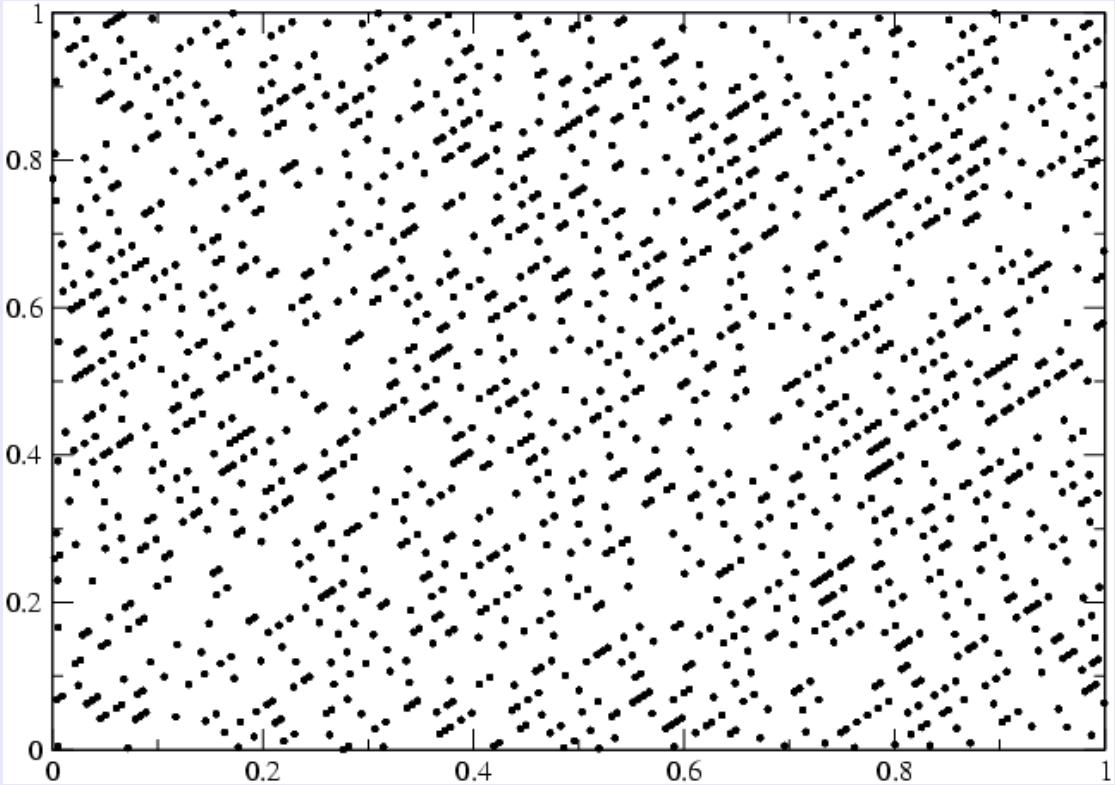
How good a generator is depends on how long we need to use it for: $\tau \ll \min(\tau_S, \tau_P, \tau_C)$

Linear Congruential Method

$$x_i = [(a \times x_{i-1} + b) \mod (c)] \qquad x_i, a, b, c \in \mathbb{N}$$
$$r = x_i / c$$

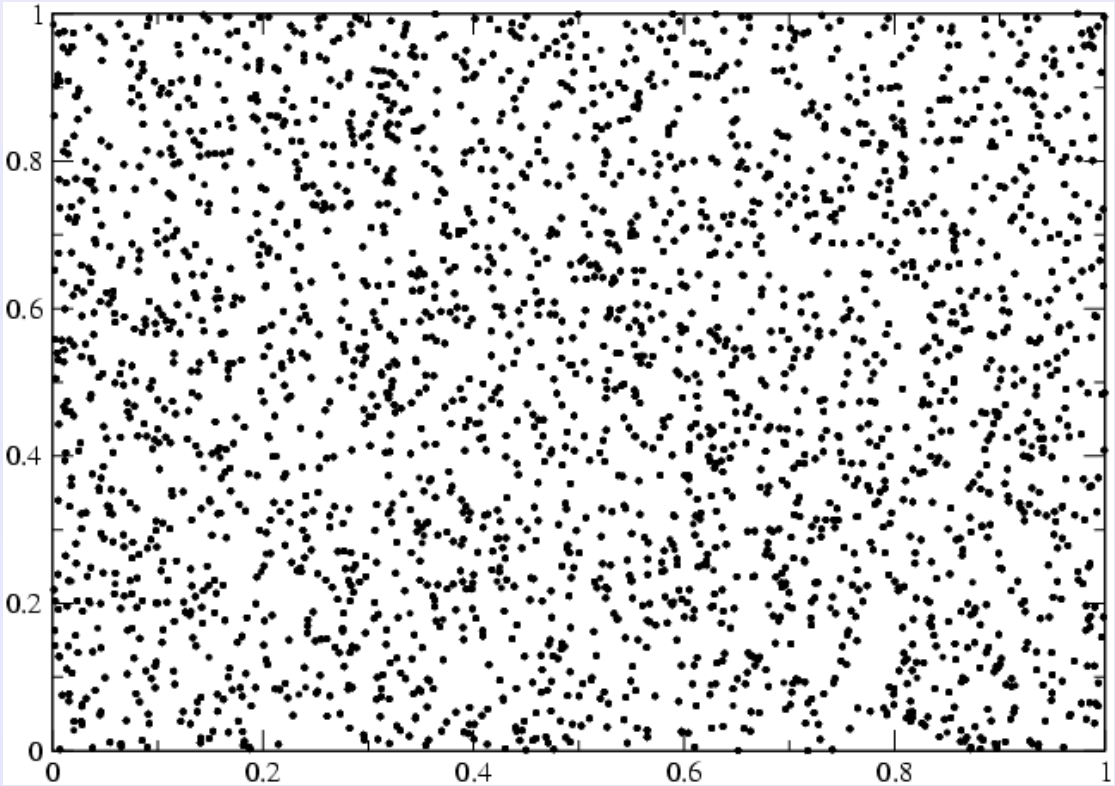
How good the generator is depends on the choice of the parameters. In particular, the period is related to c

Very bad generator



$a = 121,\ b = 0,\ c = 6133$

Good generator



$a = 135121,\ b = 0,\ c = 61331237$

Part II

Integration by Monte Carlo methods

Monte Carlo vs. Grid methods

Grid Methods

$$\text{Systematic error} \propto \mathcal{O}(1/N^{s/d})$$

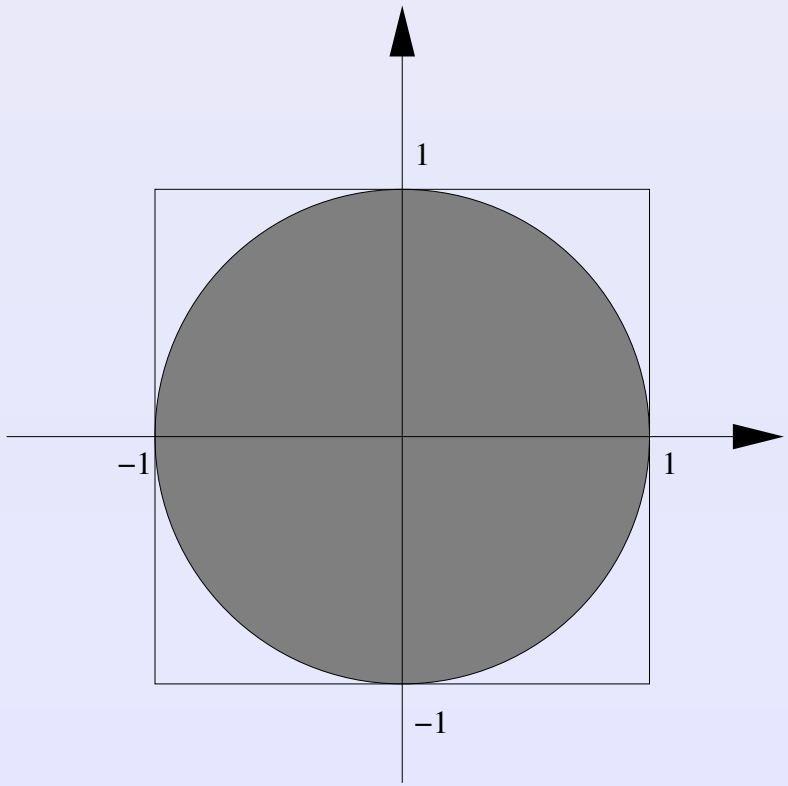
for instance, for the Simpson method $s = 4$

Monte Carlo methods

$$\text{Systematic error} \propto \mathcal{O}(1/\sqrt{N})$$

Monte Carlo methods become convenient for a large number of integration variables

Computation of π



$\pi/4 \approx$ Events inside the circle / Total number of events

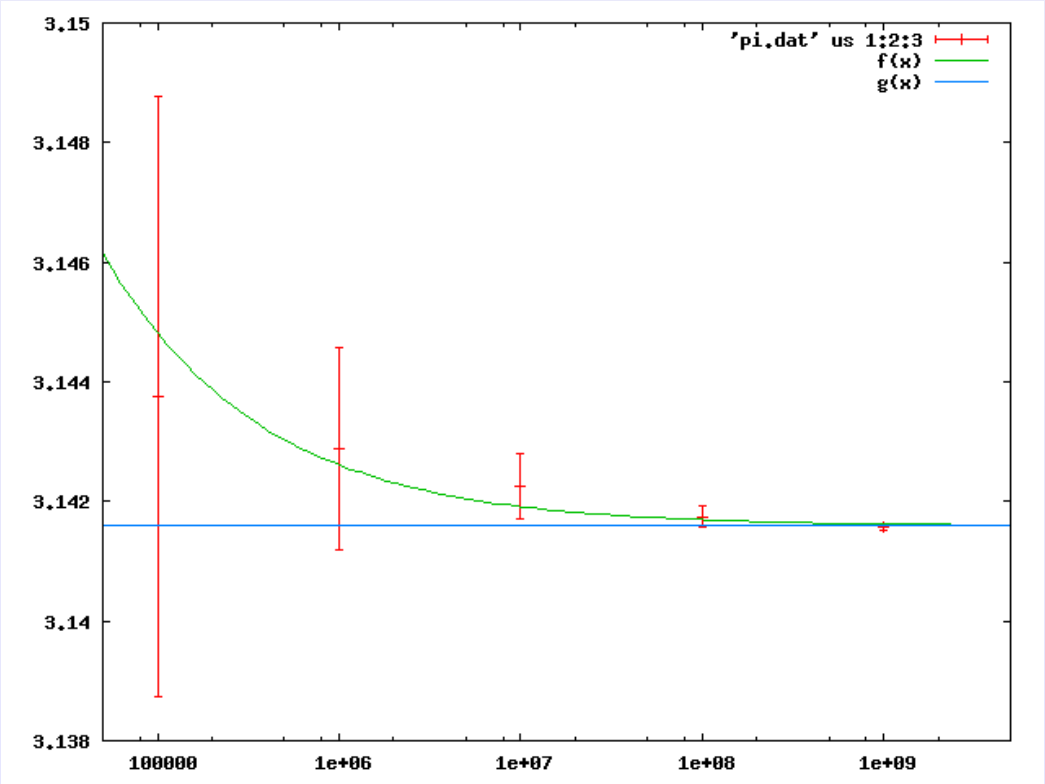
Algorithm for the computation of π

- generate pairs of random numbers in [-1;1[
- compute how many pairs fall inside the circle
- take the ratio of those over the total number of generated events

A more efficient method

$$I = \int_D F(x)dx = \frac{1}{V}\int_D F(x)Vdx = V\frac{\int_D F(x)dx}{\int_D dx} \simeq \frac{V}{N}\sum f(x_i)$$

Convergence of the estimate



Part III

Monte Carlo in Statistical Mechanics

A simple example

$$H = x^2$$

The partition function

$$Z = \int \mathrm{d}x\ e^{-\beta H}\,, \qquad \beta = 1/T$$

is exactly computable

$$P(x) = e^{-\beta H}/Z \qquad \text{probability distribution}$$

$$\langle U \rangle = \frac{1}{Z} \int \mathrm{d}x\ H(x)\ e^{-\beta H} = \frac{1}{2\beta} \qquad \text{internal energy}$$

Dynamics

Ergodic hypothesis: average over a statistical ensemble $\simeq$ average in time

Problem: give dynamics to the system

Fundamental property: at equilibrium, the configurations must follow Boltzmann distribution

Markov chains

Sequence of configurations C_m^n in which C_m^t depends only from C_n^{t-1} according to a probability distribution P_{nm} (upper index: time; lower index: number of the configuration)

Irreducible if from any C_j we can reach C_l $l > j$, i.e. if a time k exists such that $P_{jl}^k = \sum_{i_1 \dots i_n} P_{j i_1} P_{i_1 i_2} \dots P_{i_n l} \neq 0$ for any j, l

Aperiodic if $P_{ii}^k \neq 1$ for any i, k

State C_i positive if it occurs on average for a finite time

Equilibrium distribution

$\{C_i\}$ irreducible and aperiodic Markov chain with only positive states

- the equilibrium distribution exists and is unique ($\Rightarrow$ independence from the initial state)

$$\lim_{N\to\infty} P_{ij}^N = P_j$$

- the equilibrium distribution is stationary

$$P_j = \sum_i P_{ij}^1 P_i$$

- if the variance of the recurring time is finite

$$\sum_i P_i O(C_i) = \langle O \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} O(C_j)$$

Equilibrium distribution

$\{C_i\}$ irreducible and aperiodic Markov chain with only positive states

- the equilibrium distribution exists and is unique ($\Rightarrow$ independence from the initial state)

$$\lim_{N\to\infty} P_{ij}^N = P_j$$

- the equilibrium distribution is stationary

$$P_j = \sum_i P_{ij}^1 P_i$$

- if the variance of the recurring time is finite

$$\sum_i P_i O(C_i) = \langle O \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} O(C_j)$$

Equilibrium distribution

$\{C_i\}$ irreducible and aperiodic Markov chain with only positive states

- the equilibrium distribution exists and is unique ($\Rightarrow$ independence from the initial state)

$$\lim_{N\to\infty} P_{ij}^N = P_j$$

- the equilibrium distribution is stationary

$$P_j = \sum_i P_{ij}^1 P_i$$

- if the variance of the recurring time is finite

$$\sum_i P_i O(C_i) = \langle O \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} O(C_j)$$

Detailed balance

Monte Carlo dynamics: any Markovian dynamics

Problem: given a Hamiltonian, write a Markovian dynamics

Necessary condition unknown

Sufficient condition: detailed balance

$$e^{-\beta H(C_i)} P_{ij} = e^{-\beta H(C_j)} P_{ji}$$

Still freedom on the choice of P_{ij}

Metropolis Algorithm

$$P_{metro} = \min(1, e^{-\beta \Delta H}) \ , \qquad \Delta H = H_{new} - H_{old}$$

If the new configuration is not accepted, we replicate the previous one
Rejection probability minimised if

- we update one variable at a time
- the new proposed value is “close” to the previous one

A criterion for the acceptance is $N_{accepted}/N_{trial} = 0.5$

Heath Bath Algorithm

$$P_{hb} \propto e^{-\beta H_{new}}$$

The HB probability does not depend on the previous value of the variable we want to update

Compared to Metropolis

- advantage: better exploration of the configuration space
- disadvantage: requires random number with the same probability distribution as the system

Monte Carlo example

1. decide the structure of the program and the variables we need to use
2. choose an update algorithm and a starting point
3. discard the configuration needed to reach equilibrium
4. measure observables after one or more updates
5. compute the averages at the end of the program or better store observables in a file
6. use any possible trick to reduce the running time

Example

Gaussian system with dynamics

$$x_{n+1} = x_n - 2ar + a\,, \qquad r \text{ random} \in [0,1[$$

Part IV

Monte Carlo methods in Quantum Mechanics

Path Integrals in Quantum Mechanics

An alternative way to formulate Quantum Mechanics (due to Feynman) is the path integral

For a system with mass m subject to a potential $V(x)$ in addition to the Hamiltonian $H = (1/2m)p^2 + V(x)$ we define the Lagrangian $L = (1/2)m\dot{x}^2 - V(x)$

The probability amplitude of having x_{t_0} at time t_0 and x_f at time t_f is

$$\langle x_f(t)|e^{-iHt}|x_0(0)\rangle = \int (\mathcal{D}x)\, e^{iS}, \qquad S = \int_{t_0}^{t_f} d\mathrm{t}L$$

and $(\mathcal{D}x)$ is a formal expression that means "integration over all possible paths connecting x_0 and x_f

Wick rotation

The weighting with the action of the paths involves a complex integral that is not suitable for numerical computations

We perform the Wick rotation $t \to it$ and define the Euclidean version of L and S

$$L_E = \frac{1}{2} m \dot{x}^2 + V(x) \,, \qquad S_E = \int_{t_0}^{t_f} dt L_E$$

$\langle x_f(t) | e^{-iHt} | x_0(0) \rangle$ can be obtained by continuing analytically

$$\int (\mathcal{D}x)\, e^{-S_E},$$

Analogy with a statistical mechanical system with Hamiltonian S_E

Path integral and ground state

- for Z, integrate over all possible initial and final x, with the condition $x_f = x_0$
- one can prove that with this choice

$$\lim_{t_f \to \infty} Z = e^{-E_0 t_f} |c_0|^2$$

- Expectation values of observables over the ground state are given by

$$\langle O_1(t_1) \ \dots \ O_n(t_n) \rangle = Z^{-1} \int O_1(t_1) \ \dots \ O_n(t_n) \, (\mathcal{D}x) \, e^{-S_E},$$

Note the analogy with ensemble averages in Statistical Mechanics

This formulation is particularly suited for extracting information about the ground states and the first exited (see later)

Discretisation

- $\mathcal{D}x$ is a formal symbol, which needs to be defined
- one possibility is to divide the temporal extension t_f in N steps of interval a such that $Na = t_f$ (temporal lattice)
- the original theory is recovered in the limit $a \to 0 \Rightarrow$ need to choose a small a (compared to the time scale of the system)
- with this choice $\mathcal{D}x = \prod_i dx(t = ia)$, i.e. a finite but large number of integrals has to be performed $\Rightarrow$ Monte Carlo integration is a good choice

Continuum limit

- for finite a the solution is distorted by discretisation effects (see later for an example with the harmonic oscillator)
- this effect disappears in the smooth limit $a \longrightarrow 0 \Rightarrow$ need to work with small a
- remember that t_f has to be large
- a good choice takes into account these two requirements

Part V

Analysis of Monte Carlo data

Probability

Probability distribution $P(x)$

$$\bar{f} = \langle f(x) \rangle = \int f(x) P(x) dx$$

Let us define

$$\bar{x} = \langle x \rangle \qquad \text{average}$$
$$\sigma^2 = \langle (x - \bar{x})^2 \rangle = \langle x^2 \rangle - \bar{x}^2 \qquad \text{variance}$$

σ is called standard deviation

Gaussian Distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

95% of the measurements are within 2σ from the average

Central limit theorem: for $N \to \infty$ the averages of the measurements are distributed according to a Gaussian distribution with average the real value and variance σ^2/N

Bias

We have a bias when the average of the estimates does not coincide with the true value

$$\text{bias} \propto \mathcal{O}(1/N)$$

This is irrelevant in the limit of infinite measurements

It is important to remove the bias when we average non-linear functions of the measured value.

Average and Error

Estimate for the average

$$x_m = \frac{1}{N}\sum_i x_i$$

Unbiased estimator of the variance

$$\sigma_m^2 \quad = \quad = \frac{1}{N(N-1)}\sum_i (x_i - x_m)^2$$

$N < 20 \Rightarrow$ a correcting factor is needed

Thermalisation

The starting point is arbitrary

Equilibrium distribution obtained after a time τ_{eq}

Strategy: discard $n\tau_{eq}$ sweeps at the beginning

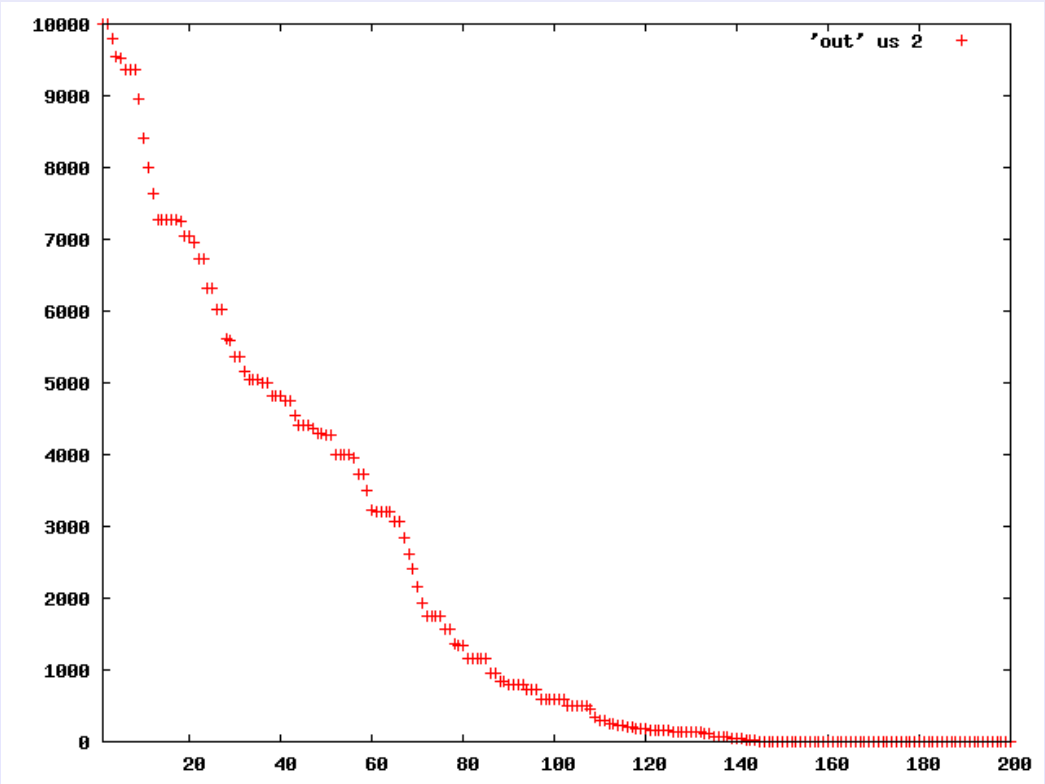
For a run with N measurements

- weight of initial sweeps $\propto n/N$
- statistical error $\propto 1/\sqrt{N}$

$\hookrightarrow$ we do not need an exact estimate of τ_{eq}

Thermalisation

Gaussian system



Common sense rules

- $\tau_{eq} \ll N$: we can be cautious and discard more sweeps than strictly needed
- $\tau_{eq} < N$: we need to estimate τ_{eq} carefully

Correlations

Hypothesis for the Gaussian analysis: uncorrelated data
The independence among the data means that only the statistical weight of the configurations determines the history of the system

However Monte Carlo dynamics limits the possibilities of moving in configuration space

Example

Dinamics $x_{n+1} = x_n - 2ar + a$

Problem: remove correlations

Correlation time

For a given observable and given dynamics

- Exponential correlation time

$$C(\tau) = \langle O(t)O(t+\tau)\rangle \propto e^{-\tau/\tau_{exp}}$$

- Integrated correlation time

$$\tau_{int} = 1 + 2\sum_{\tau=1}^{N-1} C(\tau)$$

For the error $\sigma^2 = \sigma^2_{naive}\tau_{int}$

Moreover $2\tau_{exp} \simeq \tau_{int}$

Results for the Gaussian system

Expected values for $\beta = 0.5$: $x = 0.0$ and $U = 1.0$

Naive results $x = 0.0083(33)$ and $U = 0.992(5)$

Results for x

$$\tau_{exp} = 1.90(4) \;\Rightarrow\; \langle x \rangle = 0.0083(65)$$

Results for U

$$\tau_{exp} = 1.87(3) \;\Rightarrow\; \langle U \rangle = 0.992(9)$$

For τ_{exp} the summation is truncated at $\tau = 4\tau_{exp}$

Binning

Binning: averages over groups of M consecutive data

M (often taken as 2^k) is the amplitude of the binning interval

$M \gg \tau \Rightarrow$ the partial averages are independent $\Rightarrow$ we can apply simple Gaussian analysis to them

M has to be chosen in such a way that we have at least 20 partial averages

We still have to deal with the bias

Jack-Knife method

It's a method for eliminating the bias
Jack-Knife sample $y_1, \ldots, y_N$ starting with partial averages (binned data) $y_1, \ldots, y_N$

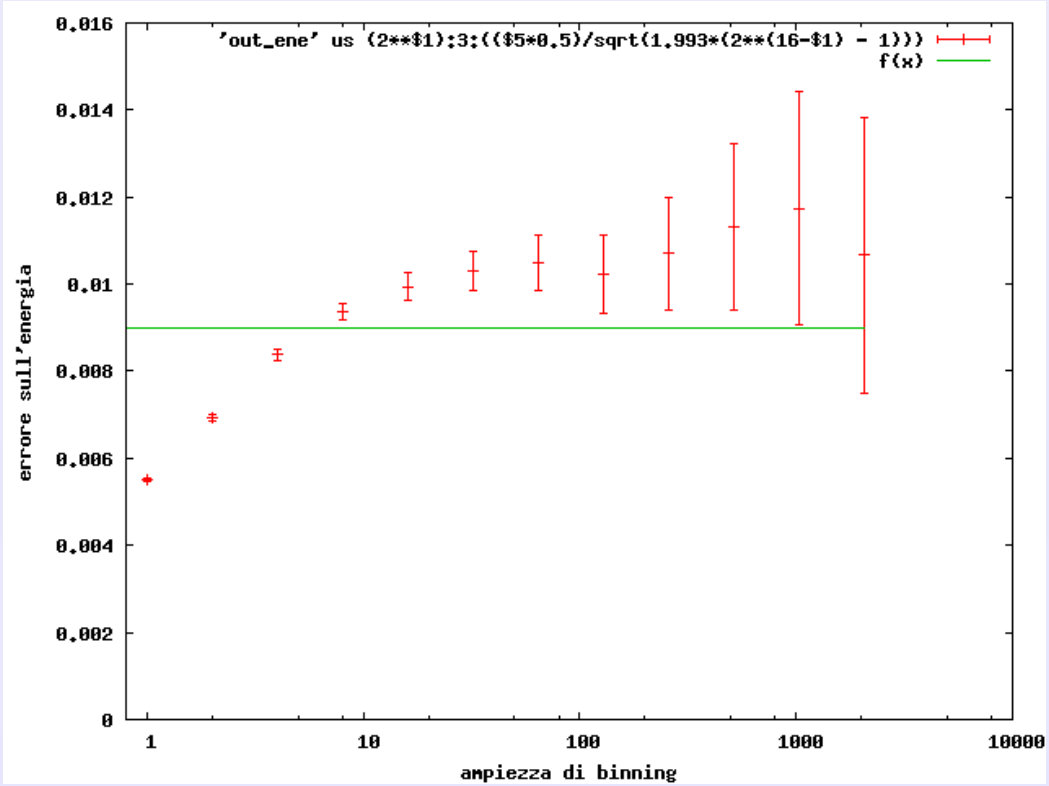
$$y_k = \frac{1}{N-1} \sum_{j \neq k} x_j$$

Elimination of the bias

$$\bar{f} = \frac{1}{N} \sum f(y_k) \qquad \text{average}$$

$$\sigma_f^2 = \frac{N-1}{N} \sum (f(y_k) - \bar{f})^2 \qquad \text{variance}$$

Jack-Knife error for the Gaussian system



One chooses the error at the plateau that has the minimal error

Fit

They are used to deduce the parameter of a theoretical behaviour from a sample of measurements

The parameters are obtained by minimising

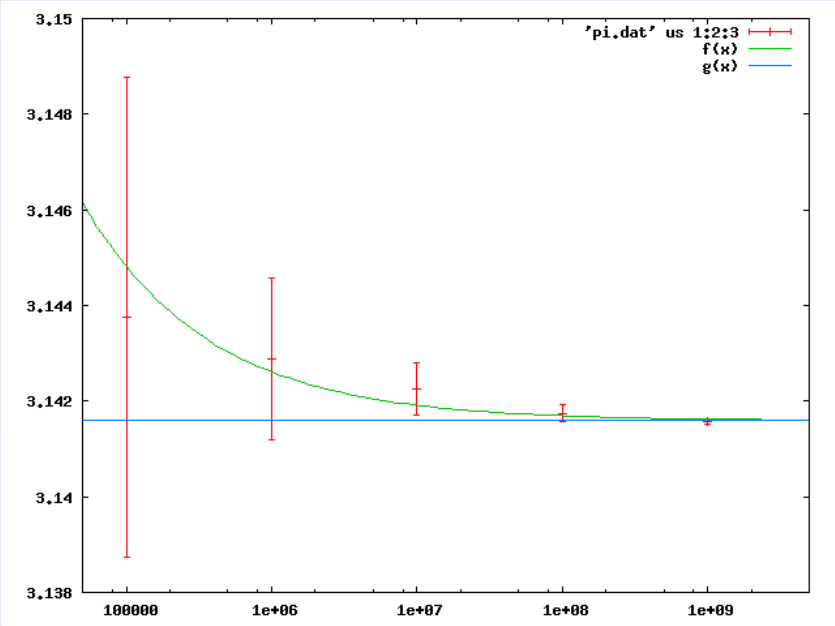
$$\chi^2(a_1, \ldots, a_m) = \sum_j \left(\frac{y^j - f(x_1^j, \ldots, x_n^j; a_1, \ldots, a_m)}{\sigma^j} \right)^2$$

Defining *dof* as the number of data minus the number of free parameters, a good fit is identified by $\chi^2/dof \simeq 1$

The error on the parameters is obtained by a Jack-knife analysis

Example: Monte Carlo error for π

We expect $|\pi - \pi(N)| = a/N^{1/2} \Rightarrow$ we want to determine a



Result $a = 1.23 \pm 0.29$ with $\chi^2/dof = 0.14$ (GNUPLOT)

Single Histogram Reweighting

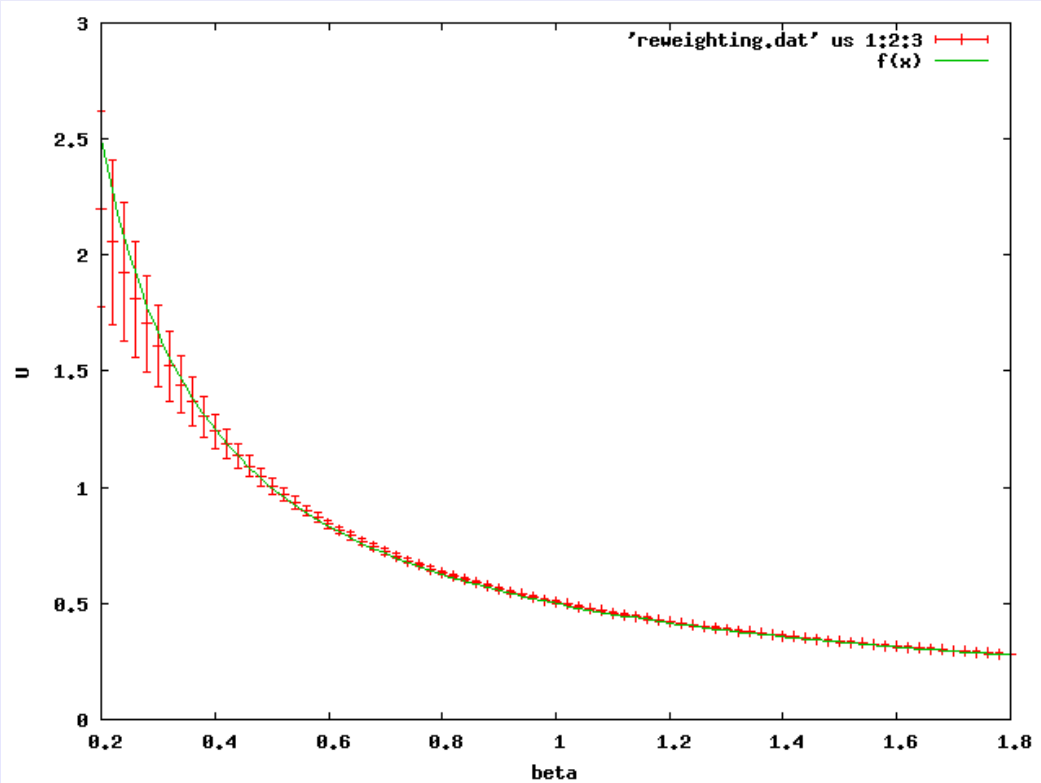
We have the following identity

$$\langle O \rangle_{\beta'} = \frac{\int dE O(E) \rho(E) e^{-(\beta'-\beta)E} e^{-\beta E}}{\int dE \rho(E) e^{-(\beta'-\beta)E} e^{-\beta E}} = \frac{\langle O e^{-\Delta\beta E} \rangle_\beta}{\langle e^{-\Delta\beta E} \rangle_\beta}$$

In principle by simulating only at one β we can obtain results for any β
In practice a Monte Carlo will never generate configurations with very low probability $\Rightarrow$ we have information only for those βs for which $\langle U \rangle_{rew}$ is less than 2σ apart from $\langle U \rangle_{orig}$

More sophisticated method: Multi Histogram

Reweighting for the Gaussian model



Starting point: $\beta = 0.9$

Monte Carlo praxis

- Identify the fundamental variables and the observables
- Write an update algorithm for the fundamental variables
- Write a measurement routine for the observables (store the values in a file for off-line analysis)
- Run the Monte Carlo with sensible parameters (thermalisation time, total number of measurements, number of sweeps between two measurements)
- Perform off-line the error analysis using binning and Jack-Knife
- If requested do some reweighting

Exercise 1: the Gaussian System

For the Gaussian system described in these notes

- write a Metropolis update algorithm
- write a routine that measures x^2
- run the Monte Carlo at $\beta = 0.1, 1, 10$ measuring x^2
- for each β reweight the results for x^2 at the other βs and comment the agreement/disagreement with the expected results

Metropolis for the Gaussian System - I

- Familiarise with the V.B. Random number generator and make sure you can extract random numbers between 0 and 1
- start with $x_0 = 1$ and define a variable α
- write an algorithm that computes

$$y = x_n - 2ar + a$$

with r random number and $a = 0.1$ to start with

- compute

$$P = e^{-\beta(y^2 - x^2)}$$

Metropolis for the Gaussian System - I

- If $P > 1$ $x_{n+1} = y$
- if $P < 1$ generate a random number r_1 in [0;1[
  1. if $r_1 < P$ $x_{n+1} = y$
  2. if $r_1 > P$ $x_{n+1} = x_n$
- generate a new y and repeat the process
- record x_{n+1} in a file The process that generate x_{n+1} from x_n (whether they are equal or not) is called "sweep" After N sweeps we generate the sequence $x_1, \ldots, x_N$

Thermalisation

To reach the statistical equilibrium the system needs a given number of sweeps N_T
The thermal equilibrium is defined by the average of the observable not changing within errors when changing the number of the sweeps
In practice, we do the following

1. discard N_1 sweeps at the beginning
2. divide the remaining ones in two samples with the same number of measurements
3. compute the average of the observables in the two samples and compare
4. if there is agreement we are at equilibrium, otherwise we discard more sweeps

Acceptance

The parameter a should be tuned in such a way that the system does not get stuck at some x_n (likely if a is big), nor moves slowly (if a is small)

This can be quantified by introducing the concept of acceptance: if $x_{n+1} \neq x_n$ our attempt at changing the variable has succeeded, otherwise it has failed.

The acceptance is defined as the ratio between the number of success over the number of sweeps; a rule of thumb is having it between 0.5 and 0.8, and we use this to choose a

Warning: the acceptance should not take into account thermalisation

Measurements

To measure an observable (e.g. x^2) we record the set of the x_N on a file and we take the simple averages

The statistical analysis should use the binning plus Jack-knife method, as described earlier, and should allow to identify τ_{exp}

To estimate the observable at another value of β we use reweighting:

$$\langle x^2 \rangle_{\beta'} = \frac{\langle x^2 e^{\Delta\beta x^2} \rangle_\beta}{\langle e^{\Delta\beta x^2} \rangle_\beta} , \qquad \Delta\beta = \beta - \beta'$$

Exercise II - The harmonic oscillator

The system described by the discretised action

$$S_E = \sum_{i=1}^{N} \left(\frac{1}{2} m (x_{i+1} - x_i)^2 + \frac{1}{2} \mu^2 x_i^2 \right)$$

with μ the elastic constant and m the mass (all variables and parameters are adimensional, i.e. pure numbers, in this problem)

Step 1: Monte Carlo Simulations

- Generalise the Metropolis algorithm for the Gaussian system to the harmonic oscillator
- How do you choose t_f and the lattice spacing? (justify)
- Run the code for $m = 1.0$ and $\mu = 1.0$ for 10000 iterations (each iteration is t_f in length) and record x at the end of each iteration

Step 2: ground state energy

The virial theorem allows to write the energy of the ground state as

$$E_0 = \mu^2 \langle x^2 \rangle$$

Based on the Monte Carlo ensemble generated at Step 1, compute E_0
Discuss carefully the errors

Step 3: reweighting

From the simulation performed, reweight the ground state energy to the points obtained by

- fixing $m = 1.0$ and taking $\mu^2 = 0.1$ and $\mu^2 = 10$
- fixing $\mu^2 = 1.0$ and taking $m = 0.1$ and $m = 10$
- taking $m = 0.2$ and $\mu^2 = 0.5$

Comment on the reliability of the results

Step 4: ground state wave function

For $m = 1$, sort the measured ensemble of the x from the smallest to the largest
Divide the interval in 100 bins of equal length and count how many data are in each bin
Assign this number to a function $f(x)$, where x is the central value of the bin, and take $\sqrt{f(x)}$ as the error on $f(x)$
Fit $f(x)$ with the formula

$$f(x) = \left(\frac{\omega}{\pi}\right)^{1/2} e^{-\omega x^2} ,$$

and compare ω with the expected value

$$\omega = \mu\sqrt{1 + \frac{\mu^2}{4}}$$

Step 5: first excited state energy

Define the quantity

$$\Delta E(\tau) = \log \frac{\langle x(0)x(\tau+1)\rangle}{\langle x(0)x(\tau)\rangle}$$

- Measure $\Delta E(\tau)$ for $\tau = 1, 2, \ldots, 10$
- Show that at large τ ΔE reaches a plateau
- Fit this plateau with the function $f(\tau) = c$, determining in this way the value of c

The energy of the first excited state can be obtained as

$$E_1 = E_0 + c$$

Exercise III - The anharmonic oscillator

Consider the harmonic oscillator with the additional term $\lambda \sum_i x_i^4$; now

$$E_0 = \mu^2 \langle x^2 \rangle + 3\lambda \langle x^4 \rangle \ ,$$

while all other formulas stay the same

- After generalising the Monte Carlo to this case, compute the energy of the ground state and of the first excited state for $m = 1.0$, $M = 1.0$ and $\lambda = 0.2,\ 1.0, 5.0$
- For the $\lambda = 1.0$, evaluate the probability distribution for x in the ground state (i.e. the square of the wave function) with a binning procedure similar to that described for the harmonic oscillator

References

1. For random numbers, Monte Carlo in Statistical Mechanics and the Gaussian System see
B.A. Berg, *Introduction to Markov Chain Monte Carlo Simulations and their statistical analysis*, arXiv:cond-mat/0410490
2. For an introduction to Path Integral methods and Monte Carlo in Quantum Mechanics see
M. Creutz, A. Freedman, *A Statistical Approach to Quantum Mechanics*, Annals Phys. 132, 427 (1981)