Monte Carlo methods

Dénes Sexty

University of Graz

2020-2021, winter term

UNIVERSITÄT GRAZ
UNIVERSITY OF GRAZ

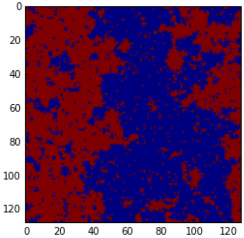
1. Examples: Ising model etc.
2. Integration vs Sampling
3. Pseudo random number generators
4. Generation of random numbers with a given given distribution
5. Percolation, Random Walks
6. Improtance Sampling, Markov chains, Metropolis Alg.
7. Statistics, error estimates: Jackknife, bootstrap
8. Langevin equation
9. fitting, χ^2 test
10. Potts model, Ising model, XY model, nonlinear $O(n)$ model
11. Monte carlo simulations in practice
12. Phase transitions, finite size effects
13. Optimization: Simulated Annealing, Genetic algorithms
14. **Cluster algorithms, Demon Algorithm**
15. **Quantum Monte Carlo**
16. **Sign Problem, Complex Langevin, Lefschetz Thimble**

Here we discuss the Ising-model.

$$H = -J \sum_{neighbors} s_i s_j - h \sum_i s_i \quad (1)$$

Near T_c there are large domains.
Spin-flip Metropolis alg. moves the boundaries of the domains
$\rightarrow$ many sweeps until a new configuration

Autocorrelation: $\tau \sim \xi^z$
$z \approx 2.125$ for 2D Ising, $z \approx 2$ for 3D Ising



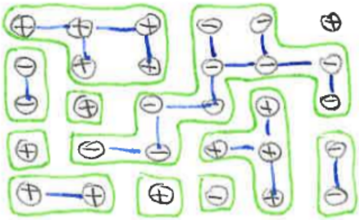
Close to T_c one a 128×128 lattice

The idea in Swendsen-Wang is to better understand the physical degrees of freedom (the domains) and try to update them.
Given a configuration

1. Visit all links which have the same spins at both ends
2. introduce a bond with probability

$$P_{bond} = 1 - e^{-2\beta J} \tag{2}$$

3. identify clusters which are connected with bonds
   (Use Hoshen-Kopelmann as discussed in connection with percolation)
4. Flip all clusters with probability $\frac{1}{2}$
5. delete all bonds → new configuration



The algorithm is **ergodic**
There is a nonzero chance that any spin becomes a cluster on its own and is flipped
$\implies$ Any configuration can be reached for any other configuration.

Detailed balance is satisfied $\implies$ we get to the correct distribution

Very effiecient As clusters can get broken up and large islands get flipped in one step.

$$\mathrm{2D}: \quad z \sim 0.33 \quad \mathrm{3D}: z \sim 0.53 \tag{3}$$

Further Advantage: Some observables measured in terms of clusters
After identification of the clusters, every cluster could be flipped with $1/2$ probability
$\Longrightarrow$ 2^{N_c} configurations can be reached, each with the same probability. We can use all of these configurations to measure some observables.

$$m = \frac{1}{V}\sum_i s_i = \frac{1}{V}\sum_{k=1}^{N_c} n_k \mathrm{sign}\, C_k \tag{4}$$

N_c is the number of clusters, n_k is the number of spins in the k-th cluster. $\mathrm{sign}(C_k)$ is the orientation in the k-th cluster.

$$\begin{aligned} \langle m^2 \rangle &= \frac{1}{V^2}\left\langle \left(\sum_{k=1}^{N_c} n_k \mathrm{sign}(C_k) \right)^2 \right\rangle = \frac{1}{V^2}\left\langle \sum_{k,k'=1}^{N_c} n_k n_{k'} \mathrm{sign}(C_k)\mathrm{sign}(C_{k'}) \right\rangle \qquad (5) \\ &= \frac{1}{V^2}\left\langle \sum_{k=1}^{N_c} n_k^2 \right\rangle + \frac{1}{V^2}\left\langle \sum_{k\neq k'} n_k n_k' \mathrm{sign}(C_k)\mathrm{sign}(C_{k'} \right\rangle \end{aligned}$$

The second term averages to zero.

also called Wolff algorithm

1. Choose a lattice point randomly (called the “seed” of the cluster)
2. Look for neighbors with the same spin and connect to them via a bond with probability $p_{bond} = 1 - e^{-2\beta J}$
3. Go to the new elements of the cluster and try their neighbors as before with probability p_{bond}
4. If the cluster's growth is finished, flip it.

No cluster identification is needed, but also no improved estimators are only partly available.
Average cluster size (p_i=prob. of choosing the seed in i-th cluster):

$$\langle n \rangle = \sum_i p_i n_i = \sum_i \frac{n_i}{V} n_i = V \langle m^2 \rangle \tag{6}$$

above T_c we have $\langle m \rangle = 0$ so we have

$$\chi = V \langle m^2 \rangle = \langle n \rangle \tag{7}$$

Generalization of the Swendsen-Wang algorithm for continous spins.

O(N) Spin model

$$H = -J \sum_{neigh.} S_i S_j, \quad S_i = \begin{pmatrix} S_i^1 \\ S_i^2 \\ \vdots \\ S_i^N \end{pmatrix}, \quad S_i S_i = 1 \tag{8}$$

S is invariant under global rotations with M orthogonal matrix

$$S_i' = M S_i, \quad M^T M = 1 \implies S_i' S_j' = S_i'^T S_j' = S_i^T M^T M S_j = S_i S_j \tag{9}$$

Spins can be decomposed in parallel and perpendicular components given a vector u with $uu = 1$

$$S_i^{\|} = (S_i u) u, \quad S_i^{\perp} = S_i - S_i^{\|} \tag{10}$$

$\implies \quad S_i^{\perp} u = S_i^{\perp} S_j^{\|} = 0$

Scalar products can be decomposed:

$$S_i S_j = (S_i^{\|} + S_i^{\perp})(S_j^{\|} + S_j^{\perp}) = S_i^{\|} S_j^{\|} + S_i^{\perp} S_j^{\perp} \tag{11}$$

we choose a random unit vector u

$$H = -J \sum_{neigh} S_i S_j = -J \sum \left[S_i^{\parallel} S_j^{\parallel} + S_i^{\perp} S_j^{\perp} \right] \tag{12}$$

We can write the parallel part as $S_i^{\parallel} = \epsilon_i |S_i^{\parallel}| u$ with $\epsilon_i = \pm 1$
We than decompose the Hamiltonian:

$$H = \underbrace{-J \sum |S_i^{\parallel}||S_j^{\parallel}| \epsilon_i \epsilon_j}_{=H^{\parallel}} \underbrace{-J \sum S_i^{\perp} S_j^{\perp}}_{=H^{\perp}} \tag{13}$$

$$H^{\parallel} = -\sum_{neigh} J_{ij} \epsilon_i \epsilon_j, \quad J_{ij} = J |S_i^{\parallel}||S_j^{\parallel}| \tag{14}$$

Effective Ising Hamiltonian with bond dependent coupling.

Now we can update ϵ_i with the Swendsen-Wang alg.

1. choose a random u vector with $uu = 1$
2. insert a bond between neighbors ϵ_i and ϵ_j with probability $p_{bond} = 1 - e^{-2\beta J_{ij}}$
3. identify clusters
4. Flip each cluster with probability $\frac{1}{2}$.

$$S_i = S_i^{\parallel} + S_i^{\perp} \;\rightarrow\; -S_i^{\parallel} + S_i^{\perp} \tag{15}$$

5. Delete all bonds

One can show that detailed balance, ergodicity is OK.

Consider a **spin chain** with spin $\frac{1}{2}$ particles.
Basis states for one spin:

$$\text{spin up: } |+1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{spin down: } |-1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{16}$$

Pauli operators:

$$\hat{\sigma}^z |s\rangle = s|s\rangle, \qquad \hat{\sigma}^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{\sigma}^x |s\rangle = |-s\rangle, \quad \hat{\sigma}^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{17}$$

We have a chain of L spins. The Hilbert space is given by:

$$|s\rangle = |s_1\rangle \otimes |s_2\rangle \otimes \ldots |s_L\rangle, \quad s_n = \pm 1, \quad \langle s|s'\rangle = \prod_{i=1}^{L} \langle s_i|s_i'\rangle \tag{18}$$

Commutation relations:

$$\begin{aligned} \forall n, m : \ [\hat{\sigma}_n^x, \hat{\sigma}_m^x] = [\hat{\sigma}_n^z, \hat{\sigma}_m^z] = 0 \\ \forall n \neq m : \ [\hat{\sigma}_n^x, \hat{\sigma}_m^z] = 0, \quad \forall n : \ [\hat{\sigma}_n^x, \hat{\sigma}_m^z] \neq 0 \end{aligned} \tag{19}$$

one dimensional Ising model in a transverse field

$$\hat{H} = \underbrace{-J\sum_{i=1}^{N}\hat{\sigma}_i^z\hat{\sigma}_{i+1}^z}_{=\hat{H}_1}\underbrace{-h\sum_{i=1}^{N}\hat{\sigma}_i^x}_{=\hat{H}_2} \tag{20}$$

We want to calculate the canonical partition function

$$Z = \mathrm{Tr}\left(e^{-\beta\hat{H}}\right) \tag{21}$$

Baker-Campbell-Hausdorf:

$$e^{(A+B)\Delta t} = e^{A\Delta t}e^{B\Delta t}e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3) \tag{22}$$

Trotter Formula

$$e^{A+B} = \lim_{n\to\infty}\left(e^{A/n}e^{B/n}\right)^n \tag{23}$$

$$e^{A/n} = 1 + \frac{A}{n} + O\left(\frac{1}{n^2}\right) \qquad e^{A/n}e^{B/n} = 1 + \frac{A+B}{n} + O\left(\frac{1}{n^2}\right)$$

$$(e^{A/n}e^{B/n})^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{A+B}{n}\right)^k + O(1/n^2)$$

$$\binom{n}{k}\frac{1}{n^k} = \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{1}{k!} = (1+O(1/n))\frac{1}{k!}$$

$$\lim_{n\to\infty}(e^{A/n}e^{B/n})^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{A+B}{k!}(1+O(1/n)) + O(1/n^2) = e^{A+B}$$

We can therefore use for small step : $e^{i(A+B)\Delta t} = e^{iA\Delta t}e^{iB\Delta t} + O(\Delta t^2)$

Higher order discretisation: $e^{i(A+B)\Delta t} = e^{iA\Delta t/2}e^{iB\Delta t}e^{iA\Delta t/2} + O(\Delta t^3)$
Exercise: Show that

$$e^{-iH_1\Delta t/2}\dots e^{-iH_L\Delta t/2}e^{-iH_L\Delta t/2}\dots e^{-iH_1\Delta t/2} = e^{-i\sum H_i\Delta t} + O(\Delta t^3)$$

$$Z_N = \mathrm{Tr}\ \left(e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2}\right)^N = \sum_s \left\langle s \left| \left(e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2}\right)^N \right| s \right\rangle =$$

$$= \sum_{s^{(0)},\ldots,s^{(N-1)}} \left\langle s^{(0)} \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s^{(1)} \right\rangle \left\langle s^{(1)} \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s^{(2)} \right\rangle \ldots$$

$$\left\langle s^{(N-1)} \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s^{(0)} \right\rangle$$

Now we calculate each matrix element:

$$\left\langle s \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s' \right\rangle = \left\langle s \left| \prod_i e^{\frac{\beta J}{N}\hat{\sigma}_i^z \hat{\sigma}_{i+1}^z} \prod_i e^{\frac{\beta h}{N}\hat{\sigma}_i^x} \right| s' \right\rangle = \prod_{i=1}^{L} e^{\frac{\beta J}{N} s_i s_{i+1}} \prod_{i=1}^{L} \left\langle s_i \left| e^{\frac{\beta h}{N}\hat{\sigma}_i^x} \right| s_i' \right\rangle$$

We can simplify the second factor using $(\hat{\sigma}^x)^2 = 1$

$$\left\langle s_i \left| e^{\frac{\beta h}{N}\hat{\sigma}_i^x} \right| s_i' \right\rangle = \left\langle s_i \left| \cosh\frac{\beta h}{N} + \sinh\frac{\beta h}{N}\hat{\sigma}_i^x \right| s_i' \right\rangle \tag{25}$$

This means:

$$\left\langle s_i \left| e^{\frac{\beta h}{N}\hat{\sigma}_i^x} \right| s_i' \right\rangle = \begin{cases} \cosh\frac{\beta h}{N} & \text{for } s_i = s_i' \\ \sinh\frac{\beta h}{N} & \text{for } s_i \neq s_i' \end{cases} = \underbrace{\sqrt{\sinh\frac{\beta h}{N}\cosh\frac{\beta h}{N}}}_{=C} \left(\frac{\sqrt{\cosh(\beta h/N)}}{\sqrt{\sinh(\beta h/N)}}\right)^{s_i s_i'} =$$

$$= C e^{\frac{\beta}{N} K s_i s_i'} \ \text{with}\ K = \frac{N}{2\beta}\ln\left(\coth\left(\frac{\beta h}{N}\right)\right)$$

$$\left\langle s \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s' \right\rangle = \prod_{i=1}^{L} e^{\frac{\beta J}{N} s_i s_{i+1}} \prod_{i=1}^{L} \left\langle s_i \left| e^{\frac{\beta h}{N}\hat{\sigma}_i^x} \right| s_i' \right\rangle = e^{\sum_{i=1}^{L} \frac{\beta J}{N} s_i s_{i+1}} \prod_{i=1}^{L} C e^{\frac{\beta}{N} K s_i s_i'} =$$
$$= C^L e^{\sum_{i=1}^{L} \frac{\beta J}{N} s_i s_{i+1} + \sum_{i=1}^{L} \frac{\beta}{N} K s_i s_i'} \quad (26$$

Now we need the product of N such matrices:

$$Z_N = \sum_{s^{(0)},\ldots,s^{(N-1)}} \prod_{t=0}^{N} \left\langle s^{(t)} \left| e^{-\frac{\beta}{N}\hat{H}_1} e^{-\frac{\beta}{N}\hat{H}_2} \right| s^{(t+1)} \right\rangle \quad (27)$$

We have periodic boundary conditions in t (and also in the i direction)

$$Z_N = C^{LN} \sum_{s^{(0)},\ldots,s^{(N-1)}} \prod_{t=0}^{N} e^{\sum_{i=1}^{L} \frac{\beta J}{N} s_i^{(t)} s_{i+1}^{(t)} + \sum_{i=1}^{L} \frac{\beta K}{N} s_i^{(t)} s_i^{(t+1)}} = \quad (28)$$
$$= C^{LN} \sum_{s_i^{(t)}=\pm 1} e^{\frac{\beta}{N} \sum_{i,t} \left[J s_i^{(t)} s_{i+1}^{(t)} + K s_i^{(t)} s_i^{(t+1)} \right]}$$

Quantum 1D Ising chain $\to$ 2D classical (anisotropic) Ising model

1. $\hat{H} = \sum \hat{H}_i$ terms in each H_i should commute with each other.
2. Trotter decomposition:

$$Z = \lim Z_N, \quad Z_N = \mathrm{Tr}\left[\left(\prod_i e^{-\frac{\beta}{N}\hat{H}_i}\right)^N\right] \tag{29}$$

3. include the unity operator: $1 = \sum_\alpha |\alpha^{(t)}\rangle\langle\alpha^{(t)}|$ between each $\prod_i e^{-\frac{\beta}{N}\hat{H}_i}$

$$Z_N = \sum_\alpha \left\langle \alpha^{(0)} \middle| \prod_i e^{-\frac{\beta}{N}\hat{H}_i} \middle| \alpha^{(1)} \right\rangle \ldots \left\langle \alpha^{(N-1)} \middle| \prod_i e^{-\frac{\beta}{N}\hat{H}_i} \middle| \alpha^{(0)} \right\rangle \tag{30}$$

4. Calculate the matrix element

$$\left\langle \alpha^{(t} \middle| \prod_i e^{-\frac{\beta}{N}\hat{H}_i} \middle| \alpha^{(t+1)} \right\rangle \tag{31}$$

To get the partition function as a sum over “configurations” $\alpha^{(t)}$

This converts a d dimensional quantum system into a $d+1$ dimensional classical statistical system, which can be dealt with using standard Monte Carlo methods.

Suppose we want to calculate in the Microcanonical ensemble (should not matter much if the system size is large)
Usually $\rightarrow$ Use the equations of motion calculated e.g. in Hamiltonian formulation $\equiv$ **Molecular dynamics**. Random sampling through chaotic behaviour. (Not discussed in detail in this lecture).

What to do with e.g. the Ising model, where no time evolution eq. is available?

Alternative: **Demon Algorithm** for calculating in the Microcanonical ensemble by Creutz. A demon travels on the lattice, and it has a bag which can contain a positive amount of energy

1. initialize a configuration with a given energy E, the bag is empty. (or lattice ground state, bag contains all energy)
2. Propose a change in the configuration
3. if the energy change is negative, the demon takes the energy (and puts it into the bag) and the change is carried out.
4. if the energy change is positive, and the demon has enough energy in its bag, it is carried out, substracting the energy from the bag.
5. otherwise the change is rejected.
6. goto 2

From the point of the demon, the system acts a big heat reservoir. E_D, the energy in the bag is distributed as:
$\implies p(E_D) \sim e^{-\beta E_D}$
This allows the measurement of a temperature.
Consider e.g. Ising model with zero magnetic field. In this case the

$$E_D = 4kJ, \quad k = 0, 1, \ldots \tag{32}$$

$$\begin{aligned} \langle E_D \rangle &= \frac{\sum_{k=0}^{\infty} 4kJe^{-\beta 4kJ}}{\sum_{k=0}^{\infty} e^{-\beta 4kJ}} = -\frac{\partial}{\partial\beta} \ln \sum e^{-\beta 4kJ} = -\frac{\partial}{\partial\beta} \ln \frac{1}{1-e^{-\beta 4J}} \\ &= \frac{4Je^{-\beta 4J}}{1-e^{-\beta 4J}} = \frac{4J}{e^{4\beta J}-1} \end{aligned} \tag{33}$$

This implies $\beta = \frac{1}{4J} \ln(1 + 4J/\langle E_D \rangle)$.
For continous models we have

$$\langle E_D \rangle = \frac{\int_0^{\infty} E_D e^{-\beta E_D} dE_D}{\int_0^{\infty} e^{-\beta E_D} dE_D} = -\frac{\partial}{\partial\beta} \ln \int e^{-\beta E_D} dE_D = -\frac{\partial}{\partial\beta} \ln \frac{1}{\beta} = \frac{1}{\beta}$$

We are interested in a system described with

$$Z = \int d\Phi e^{-S} = \mathrm{Tr}(e^{-\beta(\hat{H}-\mu\hat{N})}) = \sum_C W(c) \tag{34}$$

We have learned about improtance sampling so far:
We build a Markov chain of configurations (Metropolis algorithm, Langevin eq., etc.)

$$\ldots \to C_{i-1} \to C_i \to C_{i+1} \ldots \tag{35}$$

where we arranged the properties of the chain such that $p(C) \sim W(C)$ (probability of visiting C proportional to the weight)

$$\langle X \rangle = \frac{1}{Z}\mathrm{Tr}Xe^{-\beta(\hat{H}-\mu N)} = \frac{1}{Z}\sum_C W(C)X(C) = \frac{1}{N}\sum X[C_i] \tag{36}$$

If we have $W[C]$ non positive (or even complex) this strategy breaks down. In that case we have a **Sign Problem**.

Sometimes we can solve the Sign problem by changeing the representation:

$$Z = \sum_C W[C] = \sum_D W'[D] \tag{37}$$

if $Z \geq 0$ we might be able to find a representation with positive terms $W'[D] \geq 0$. We can than simulate in terms of $D \implies$ We need to find the right variables.

Sometimes Z is non positive, in that case we have a sign problem in every representation (e.g. at complex parameters, in supersymmetry, etc.)

Time evolution operator: $U = e^{-it\hat{H}}$. e.g.: $|\Psi(x,t)\rangle = e^{-it\hat{H}}|\Psi(x,t=0)\rangle$.
We are interested in the transition amplitude

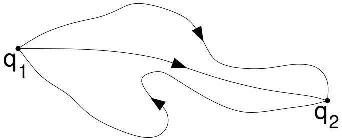
$$\langle q_2|e^{-it\hat{H}}|q_1\rangle \tag{38}$$

We can calculate e.g. using the Schrödinger eq.: $i\partial_t\Psi = \hat{H}\Psi$.

Equivalently: **Path integral formulation**

$$\langle q_2|e^{-it\hat{H}}|q_1\rangle = \int_{q_1}^{q_2} Dq e^{iS[q(t)]} \tag{39}$$

path integral is the sum for all functions $q(t)$ with the correct boundary conditions $q(t=t_1)=q_1, \ q(t=t_2)=q_2$



$$\langle q_2 | e^{-it\hat{H}} | q_1 \rangle = \int_{q_1}^{q_2} Dq e^{iS[q(t)]} \tag{40}$$

if we replace: $it \to \beta$, and do a Trace, we get thermal physics. (this trick is called imaginary time formalism)
If we want to calculate time evolution $\to$ Hard sign problem.

We sometimes have a sign problem if we introduce fermions. Famous example: Imbalanced Fermi gas

We often have sign problem if we introduce chemical potential. e.g.: Bose Gas, XY model, QCD, etc.

Topological terms in gauge theories also lead to a sign problem

Sometimes you can solve exactly using **new variables**

$$Z = \sum_C W[C] = \sum_D W'[D] \tag{41}$$

or using **subsets**

$$Z = \sum_C W[C] = \sum_S \left(\sum_{C \in S} W[C] \right) \tag{42}$$

If you find a good parameter for that, you can use **Taylor expansion**

$$Z(\mu) = Z(\mu = 0) + \mu \partial_\mu Z(0) + \frac{\mu^2}{2} \partial_\mu^2 Z(0) + \ldots \tag{43}$$

For this one needs to have no sign problem at $\mu = 0$.
coefficients are observables such charge density end susceptibility

$$n = \frac{1}{VZ} \partial_\mu Z(\mu), \quad \chi_q = \frac{1}{VZ} \partial_\mu^2 Z(\mu) \tag{44}$$

We want to calculate

$$\langle X\rangle = \frac{\sum_c W[C]X[C]}{\sum_c W[C]} \tag{45}$$

Suppose we come up with a modification of the weight $W[C] \to W'[C]$ such that $W'[C] > 0$. (We can use e.g. $W' = |W|$)

$$\begin{aligned}\langle X\rangle_W &= \frac{\sum_c W[C]X[C]}{\sum_c W[C]} = \frac{\sum_c W'[C](W[C]/W'[C])X[C]}{\sum_c W'[C](W[C]/W'[C])} \\ &= \frac{\frac{1}{Z'}\sum_c W'[C](W[C]/W'[C])X[C]}{\frac{1}{Z'}\sum_c W'[C](W[C]/W'[C])} = \frac{\langle (W/W')X\rangle_{W'}}{\langle (W/W')\rangle_{W'}}\end{aligned} \tag{46}$$

Where we defined $Z' = \sum_c W'[C]$

if we have $W' = |W|$ than $W/W' = e^{i\theta}$.
For $\langle W/W'\rangle_{W'} \sim O(1)$ we have a mild sign problem,
For $\langle W/W'\rangle_{W'} \ll 1$ we have a severe sign problem (and this method fails).

Let's look again at $\langle (W/W') \rangle_{W'}$

$$\langle (W/W') \rangle_{W'} = \frac{\sum_C W'(W/W')}{\sum_C W'} = \frac{\sum_C W[C]}{\sum_C W'[C]} = \frac{Z_W}{Z_{W'}} \tag{47}$$

Using the free energy: $F = -k_B T \ln Z$

$$\langle (W/W') \rangle_{W'} = \frac{Z_W}{Z_{W'}} = e^{-\beta F_W} / e^{\beta F_{W'}} = e^{-\beta V \Delta f} \tag{48}$$

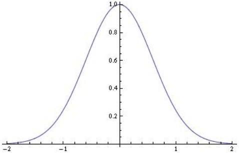
Where Δf is the difference of the free energy density between the two ensembles (F is extensive).

$\Longrightarrow$ Sign problem is exponentially hard with the volume (and usually we want to have $V \to \infty$).

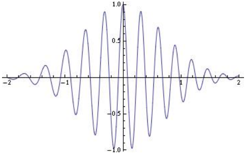
$$Z = \int_{-\infty}^{\infty} e^{-(\sigma x^2 + i\lambda x)} dx, \quad \langle x^2 \rangle = \frac{1}{Z} \int_{-\infty}^{\infty} x^2 e^{-(\sigma x^2 + i\lambda x)} dx = ? \tag{49}$$

We use random uniform sampling in the region $-a \leq x \leq a$ to estimate integrals

$$\int_{-a}^{a} f(x) dx \approx \frac{1}{N} \sum_i f(x_i) \tag{50}$$



$\sigma = \sqrt{2}, \ \lambda = 0$
We need 100 samples to get 10% relative error



$\sigma = 1 + i, \ \lambda = 20$
$Z \approx 10^{-22}$
$\sim 10^{46}$ samples for 10% relative error

We used Langevin equation before:

$$\frac{dx_i}{d\tau} = -\frac{\partial S}{\partial x_i} + \eta_i \tag{51}$$

We never had to mention probabilities $\to$ use it for a complex action.
x_i becomes complex
Observables are calculated using complex continuation (with $x \to x + iy$)

$$\langle O[x] \rangle = \frac{1}{T} \int d\tau O[x(\tau)] \ \to \frac{1}{T} \int d\tau O[x(\tau) + iy(\tau)] \tag{52}$$

For example:

$$\langle x^2 \rangle \to \langle x^2 - y^2 \rangle + i2 \langle xy \rangle \tag{53}$$

Using the complex measure $\rho(x) = \frac{1}{Z} e^{-S(x)}$ and the real probability measure $P(x, y)$ on the complex plane this means

$$\int dx \rho(x) O(x) \ \to \ \int dx dy P(x, y) O(x + iy) \tag{54}$$

$$Z = \int_{-\infty}^{\infty} e^{-(\sigma x^2 + i\lambda x)} dx, \quad S = \sigma x^2 + i\lambda x \tag{55}$$

Now we complexify the Langevin equation $S(x) \to S(z)$, and $x \to z = x + iy$

$$\frac{\partial S(z)}{\partial z} = 2\sigma z + i\lambda \tag{56}$$

One can also calculate the derivate first and complexify afterwards (since $S(x)$ is analytic, we get the same).

$$\begin{aligned} \frac{dx}{d\tau} &= -\frac{\partial S}{\partial z_i} + \eta_i = -2\mathrm{Re}(\sigma(x+iy)) - \mathrm{Re}(i\lambda) + \eta_i \\ \frac{dy}{d\tau} &= -2\mathrm{Im}(\sigma(x+iy)) - \mathrm{Im}(i\lambda) \end{aligned} \tag{57}$$

To measure the original $\langle x^2 \rangle$ in the complexified theory we measure $x^2 - y^2$.

Gaussian Example

$$S[x]=\sigma x^2+i\lambda x$$

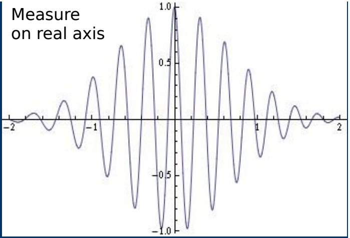
CLE

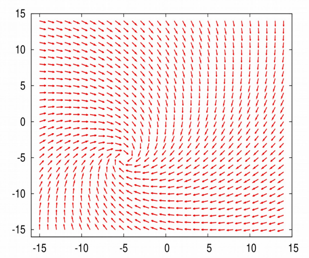
$$\frac{d}{d\tau}(x+iy)=-2\sigma(x+iy)-i\lambda+\eta$$

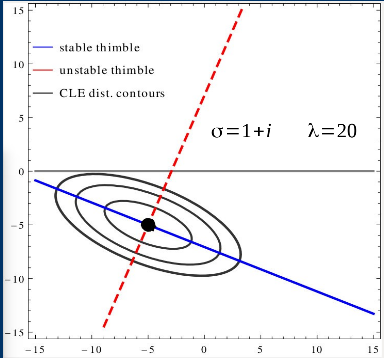
$$P(x,y)=e^{-a(x-x_0)^2-b(y-y_0)^2-c(x-x_0)(y-y_0)}$$

Gaussian distribution around critical point

$$\left.\frac{\partial S(z)}{\partial z}\right|_{z_0}=0$$







Sometimes Complex Langevin gives a spectacular solution
Other times it converges to a wrong result: process wanders to far, fluctuations grow large

When does it give a good solution?
1. Action needs to be analytical (also no poles)
2. $P(x, y)$ needs to vanish fast enough as $x, y \to \infty$.

Large amount of freedom, reparametrizations, kernels, etc.

We want to calculate the integral

$$\int_{-\infty}^{\infty} F(x)e^{-S(x)}dx \tag{58}$$

If $S(x)$ and $F(x)$ is analytic, we consider them as complex functions $S(z)$ and $F(z)$. Assuming they have no poles:

$$\int_{-\infty}^{\infty} F(x)e^{-S(x)}dx = \int_{C} F(z)e^{-S(z)}dz \tag{59}$$

Where the C curve goes from $-\infty$ on the real axis to ∞ on the real axis. It can take an arbitrary shape in between. C is parametrized as $z(t)$

$$\int_{-\infty}^{\infty} F(x)e^{-S(x)}dx = \int dt \left(\frac{dz}{dt}\right) e^{S(z(t))}F(z(t)) \tag{60}$$

it is easier to simulate on the curve C if $e^{\mathrm{Re}S(z(t))}$ has a sharp peak, and $e^{i\mathrm{Im}S(z(t))}$ is a mild sign problem

Constant imaginary part $\Longrightarrow$ no sign problem

$$\frac{d\mathrm{Im}S}{dt} = \frac{\partial ImS}{\partial x}\frac{dx}{dt} + \frac{\partial ImS}{\partial y}\frac{dy}{dt} = 0 \tag{61}$$

$$\frac{\partial ImS/\partial x}{dy/dt} = -\frac{\partial ImS/\partial y}{dx/dt} \underbrace{\Longrightarrow}_{Cauchy-Riemann} \frac{\partial ReS/\partial y}{dy/dt} = \frac{\partial ReS/\partial x}{dx/dt} \tag{62}$$

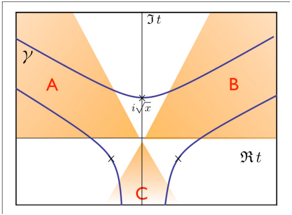
Which means the curve will be in the direction of the gradient of the real part of the action $\Longrightarrow$ sharply peaked.

Mild sign problem $\Longleftrightarrow$ sharply peaked action
This curve is called the **Lefschetz thimble**

generally we have multiple contributing thimbles

$$Z = \sum_k m_k e^{-i\mathrm{Im}S(z_k)} \int_{C_k} dt \left(\frac{dz_k}{dt}\right) e^{-\mathrm{Re}S(z_k(t))} \tag{63}$$

m_k is an integer, the intersection number
The Jacobians give a residual sign problem if the thimble is curved



In practice in makes sense to map the real axis somewhere close to the thimbles, but not necceseraly exactly on them: **Sign optimized manifolds**