# LGT4HEP Class 2

Q.) How to sample arbitrary distribution?

* Input is a sequence of random no.s uniformly distributed b/w 0–1.
* Exactly
* function of a random variable is also a random variable
* We define a probability distribution function PDF for continuous case,

$$P_X(x) - \underline{\underline{\text{PDF}}} \quad \text{in } x \in [a,b]$$

$$P(X \leq x) \equiv F_X(x) = \int_a^x dx \; P_X(x)$$

Random Value label

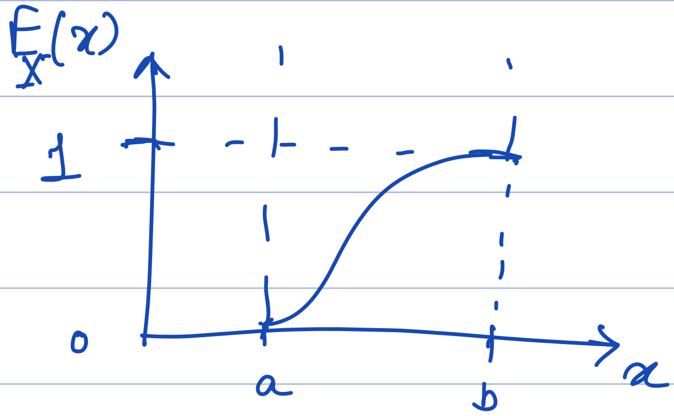
$- \underline{\underline{\text{CDF}}} -$ Probability that $X$ is lower than a chosen $x$.

Q) How can we sample PDF $P_X(x)$?

Q) What is $F_X(X)$?

→ monotonic fn of $x$

→ PDFs are normalized such that the total area is 1

→ 


Prob at upper bound is 1.

we call $Y = F_X(X)$ is a random variable

Q) How is $Y$ distributed?

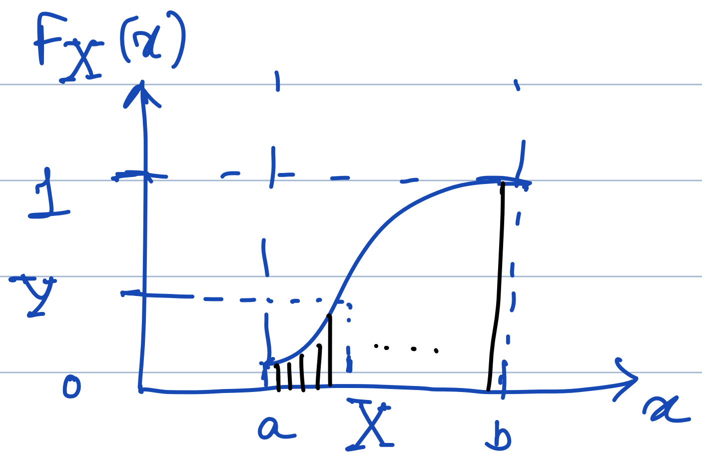
i.e. What is this function $F_Y(y) = ?$

By definition, $F_Y(y) = P(Y \leq y)$

$Y, y \in [0, 1]$

→ switch to discrete selection of intervals

$y = F_X(x)$

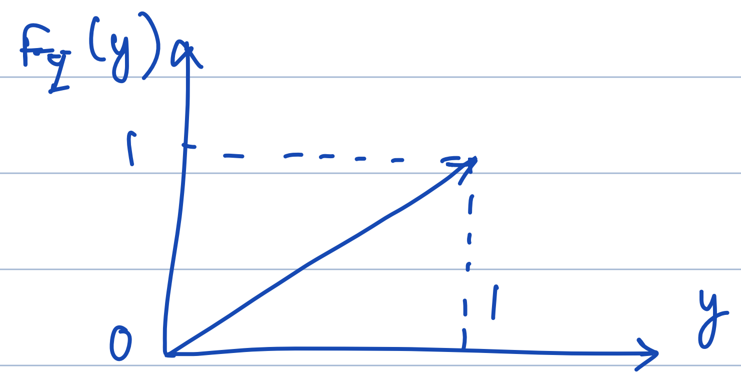



$F_Y(y) = P(Y \leq y) = P(X \leq x) = F_X(x) = y$

↳ sequence

The CDF for variable y

$\therefore \quad F_Y(y) = y$



∴ we know that:-

$$P_Y(y) = \frac{dF_Y(y)}{dy} = 1$$

→ The random variable $Y$ is uniformly distributed.

⇒ we now want to map backwards,
$X$ is what I want to generate :

$$\boxed{X = F_X^{-1}(Y)}$$

⇒ $Y$ is uniform in $[0,1]$
$X$ is from & distributed as $P_X(x)$

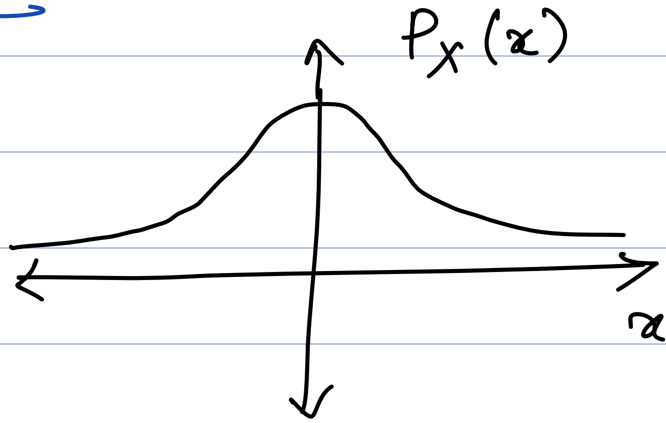
→ Some cases, $F_X(Y)$ is invertible.
it is easy to turn sequence of uniform random variables in 0-1 range into a distributed set of random no. 1.

These are prerequisites to more complicated distributions.

This is called $\boxed{\text{CDF Inversion}}$ method.

eg. Cauchy Distribution:

$$P_X(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$$




$x \in (-\infty, \infty)$

$$\Rightarrow y = F_X(x) = \int_{-\infty}^{x} dx' \, \frac{1}{\pi} \cdot \frac{1}{1+x'^2}$$

$$= \frac{1}{\pi} \arctan(x') \Big|_{-\infty}^{x'}$$

$$= \frac{1}{\pi} \arctan(x) + \frac{1}{2}$$

---

$$\Rightarrow X = \tan\left(\pi Y - \frac{\pi}{2}\right), \quad Y \in [0, 1]$$

Q) How to sample more complex Distributions?

$$\tilde{p}(x) \sim (1-x^2)^{3/2} \quad , \quad x \in [-1,1]$$

$\rightarrow$ integrable but not trivial.

$$\int_{-1}^{x} \tilde{p}(x')\,dx' = \frac{1}{8}\left[x(5-2x^2)\sqrt{1-x^2} + 3\arcsin(x) + \frac{3\pi}{2}\right]$$

$\rightarrow$ usually worse than this

$\rightarrow$ Normalized distribution

$$p(x) = \frac{8}{3\pi}(1-x^2)^{3/2}$$

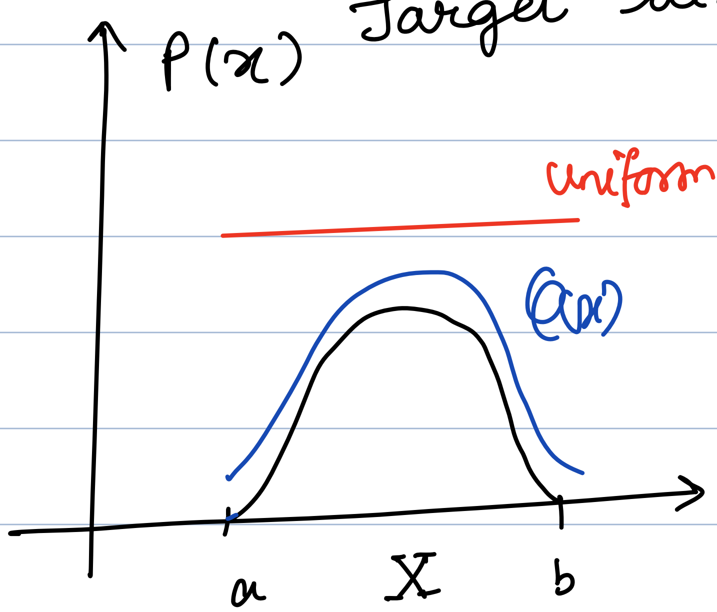
→ It is not so easy to invert this CDF can be done but will have to solve the N.L. $eq^n$ every time we produce a variable.

→ Practical way around is to create a table around it, but impractical

→ Some other way ?

Rejection method : Heat Bath

Target distribution




→ we have some prob. density defined in an interval.

→ Generate a random variable distributed in this way:

→ $P(x)$ = Target Dist.

→ Try to find an auxillary / proposal distribution. $Q(x)$
(which can be differently normalized.)

→ We need a proposal distribution such that CDF can be inverted.

The algorithm :-

① Draw $X$ from $Q(x)$

② Draw a uniform $r \in [0,1]$ (random number)

③ If $r \leq P(x)/Q(x)$, take $X$, otherwise repeat from step ①.

→ Loop breaks only when we generated a new value.

→ We want to check the following probability for a continuous distributions.

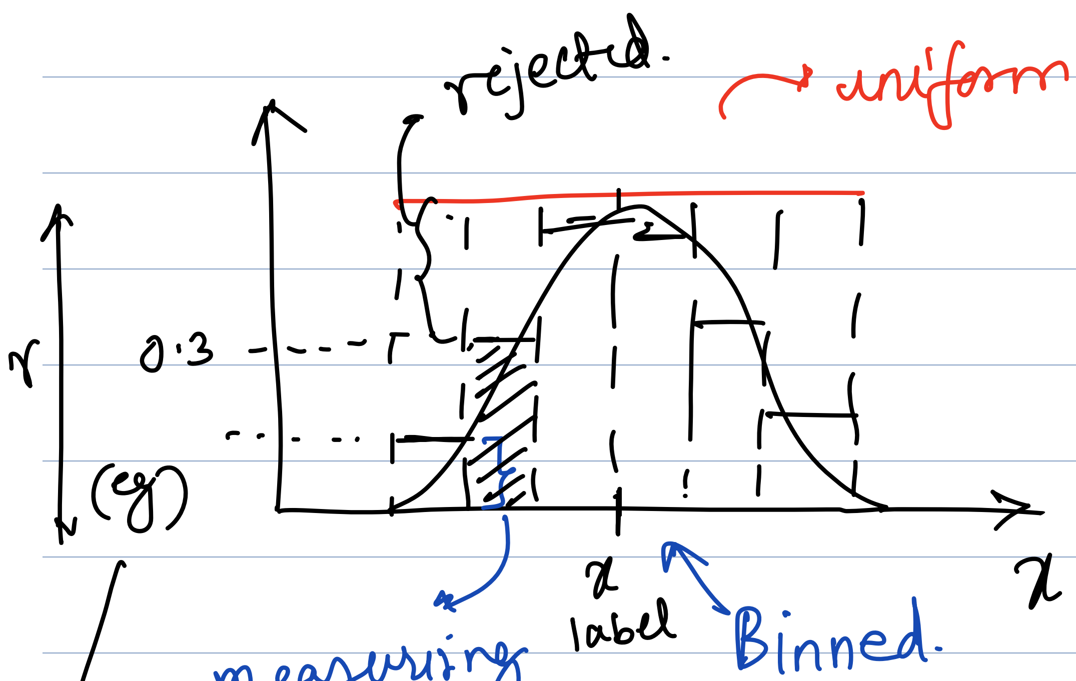
$$P(X \le x)$$

$$= \underset{x' \le x}{\sum^{\text{"or"}}} \underbrace{\overbrace{\varphi_X(x')}^{\text{p. Density}} \cdot \overbrace{\Delta x'}^{\text{width}}}_{\text{Probability of } X \text{ to be in some bin around point } x.} \overset{\text{"and"}}{\times} \underbrace{\frac{P(x')}{\varphi(x')}}_{\text{probability that the no. is generated \& accepted.}}$$

This will generate random numbers from $\varphi$.

$x' \to$ running label

$x \to$ label




Now, everything is binned.

→ Randomly deciding whether to take this number or not, but it is a probability

$\frac{p(x)}{g(x)} = 0.3$, then only 30% cases will be accepted.

Accept everything → Uniform distribution

Probability that condition 3 is satisfied is this ratio : $p(x')/g(x')$

→ I need to hit every bin to the left of '$x$', by generating variable from $\varphi(x)$, and accepted repeat until

"and"

$$p(x')/\varphi(x') \xrightarrow[\Delta x \to 0]{} \int_a^x p(x')\,dx$$

Target Prob. density.

→ Algorithm indeed samples the target distribution $p(x)$.

→ Efficiency of the method is the area under the distribution

Ask

closer are $\varphi$ to $p$, more efficient the method.

$$\boxed{\text{efficiency} : \frac{\int \varphi(x)\,dx}{\int p(x)\,dx}}$$

① ② ③ we have to repeat until the condition is satisfied.

→ This method would not scale up to a multidimensional distribution

→ Integration → Inversion (analytically)

$\underbrace{\phantom{Integration \to Inversion}}_{\text{Ridiculous for 10 dim.}}$

→ But, Separable high dimentional distribution can be used.

→ For fermions, we need more sophesticated algorithms. This is still very powerful.

METROPOLIS ALGORITHM :
- generic feature
- very imp. idea.

↳ used as a solⁿ to obtain the transition matrix T.

. Similar to rejection method.

① choose trial $\tilde{x}'$ (being at $x$)

from PDF called $g(x')$ or

generically $g(x', x)$.

eg - spin configuration where only 1 change is seen, $x'$ is changed, but $g$ will realistically also have some dependence on the surrounding.

If $g(x', x)$ has no $x'$ dependence, it produces independent $x'$ (best case scenario)

So, $g$ has two arguments, the trial state and the previous state.

So, we have a trial configuration.

Then,

(2) → Accept $x'$ $(x' = \tilde{x})$ with probability :-

$$T_A(\tilde{x}', x) = \min\left\{1, \frac{P(\tilde{x}')}{Q(\tilde{x}', x)} \Big/ \frac{P(x)}{Q(x, \tilde{x}')}\right\}$$

↓

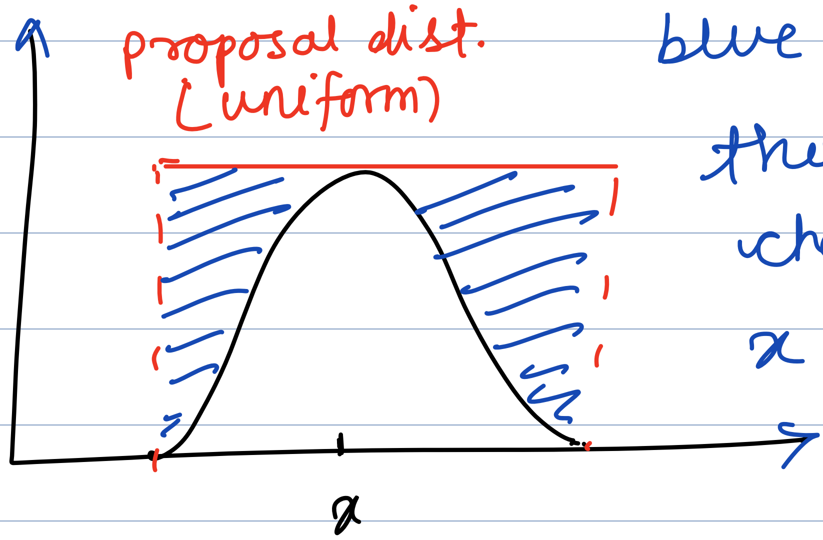
Defining feature of the Metropolis Algo. is that the acceptance prob. is a fn. of what we had before and what we are proposing.

$T_A(\tilde{x}', x)$ should satisfy the Detailed Balance equation.

(3) → If rejected, take $x' = x$. Record the state we were in again as a new state.

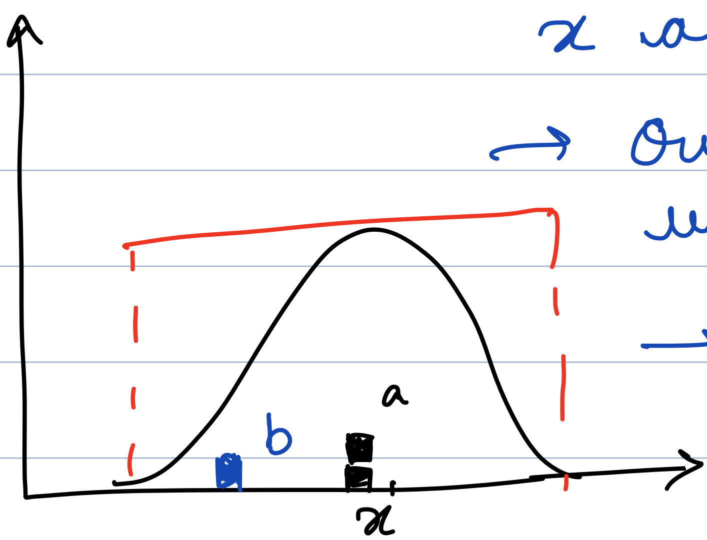
If we think abt <u>Rejection method</u>

first :




→ carring out the blue part, so there is a higher chance to be near x than the edges

Now, <u>Metropolis</u> is doing this in an entirely different way.




→ Let's say I aded up near x and put a break (a)

→ Our proposal dist. is uniform again

→ choose some 'different' trial configuration (b) but the probablity to be in the old coefficient (a) is much higher than to move some where else so put 2 bricks on (a).

→ Histogram will be populated / built more where the probability is higher.

→ and after multiple iterations, we get the full distribution of Histogram.

- Total transition Probability

$$= \left(\text{Prob to generate } \tilde{x}'\right) \times \left(\text{Prob. to make } T_A(\tilde{x}', x) \text{ accept}\right)$$

"and"

$$T(x \to x') = Q(x', x)\, T_A(x', x)$$

We want to check that:

$$\underbrace{T(x \to x')}_{\text{Prom of transitioning to } x'}\ \underbrace{P(x)}_{\text{Prob. of being at } x} = P(x)\, Q(x', x) \cdot \min\left\{1, \frac{P(x')}{Q(x', x)} \bigg/ \frac{P(x)}{Q(x, x')}\right\}$$

$$= \min\left\{P(x)Q(x', x),\ P(x')\, Q(x, x')\right\}$$

$$= \min\left\{ \frac{P(x)Q(x', x)}{P(x')Q(x, x')}, 1 \right\} \cdot P(x')Q(x, x')$$

$$= T(x' \to x) \cdot P(x')$$

RHS of detailed balance $eq^n$.

At this early we assume that Q is symm.

If Q is symmetric,

$$Q(x, x') = Q(x', x)$$ , but in general,

$$Q(x, x') = Q(x', x) \sim 1$$

$$\Rightarrow T_A = \min\left\{ 1, \frac{P(x')}{P(x)} \right\}$$

$$P(x) \sim \exp\{-\beta E(\sigma)\}$$

spin configuration

$$\Rightarrow T_A = \min\left\{ 1, \exp\left\{ -\beta\left( E[\sigma'] - E[\sigma] \right) \right\} \right\}$$

If $E[\sigma'] < E(\sigma)$, $\Rightarrow$ always Accepted

If the new conf. ends up decreasing energy, it's always accepted.

But, $E[\sigma'] > E[\sigma]$, $\Rightarrow$ Conditionally accepted.

A real distinctive feature of the Metropolis algorithm different from CDF Inversion / Rejection method is that we have to know some normalization for the later 2.

$$P(x) = \frac{1}{Z} \cdot e^{-\beta E[\sigma]}, \quad Z = \sum_{\{\sigma\}} e^{-\beta E[\sigma]}$$

partition fn.

→ Because of the way transition prob. is defined, Overall normalization cancels out, so it's enough to have the relative weight and the term.

→ It is guaranteed to converge/equilibrate to the stationary distribution.

→ This is a powerful feature as I don't need to know the overall normalization to be able to sample that distribution.

→ For any realistic system, the overall normalization is a non-trivially hard to get, it's the partition fn. calculated.

→ For a decent acceptance rate, one would need some proposal distribution which is either with the CDF Inversion method / some other method where we can control and produce things quickly.

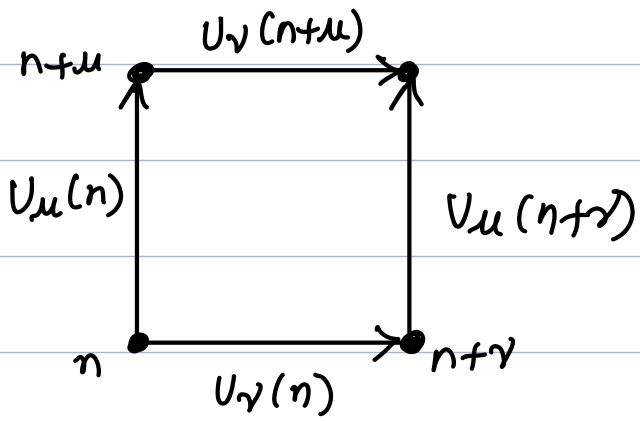
# $SU(N_c)$ Pure Gauge Theory

→ The Path Integral on the lattice

$$Z = \int \mathcal{D}[U]\, e^{-S_G[U]},$$

$$S_G[U] = \frac{\beta}{N_c} \sum_{n\in\Lambda} \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{Tr}\left\{ \mathbb{1} - U_{\mu\nu}(n) \right\}$$

→ Plaquette: $U_{\mu\nu}(n) = U_\mu(n) U_\nu(n+\hat{\mu})\, U_\mu^\dagger(n+\hat{\nu})\, U_\nu^\dagger(n)$




## Staples:

→ consider a link $U_\mu(n)$ and what plaquettes it contributes to :-

$U_{\mu\nu}(n)$ , $U_{\mu\nu}(n-\nu)$ for any $\nu \neq \mu$

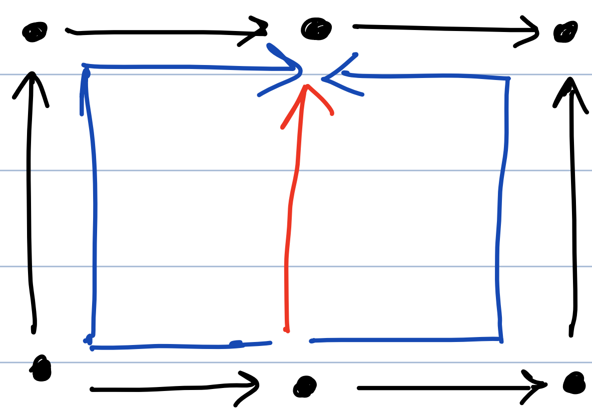
→ Rewriting contribution to the action differently :

$$\mathrm{Re\,Tr}\{U_{\mu\nu}(n)\} = \mathrm{Re\,Tr}\{S^{\dagger}_{\mu,+\nu}(n)\,U_{\mu}(n)\}$$

$$\mathrm{Re\,Tr}\{U_{\mu\nu}(n-\nu)\} = \mathrm{Re\,Tr}\{S^{\dagger}_{\mu,-\nu}(n)\,U_{\mu}(n)\}$$

$$S_{\mu,+\nu}(n) = U_{\nu}(n)\,U_{\mu}(n+\nu)\,U_{\nu}^{\dagger}(n+\mu)$$
$$S_{\mu,-\nu}(n) = U_{\nu}^{\dagger}(n-\nu)\,U_{\mu}(n-\nu)\,U_{\nu}(n-\nu+\mu)$$

→



→ SU(2) gauge Theory.

→ The contribution to the action for $U_{\mu}(n)$:

$$S_g[U_\mu(n)] \sim -\frac{\beta}{N_c} \operatorname{Re} \operatorname{Tr}\left\{ \tilde{S}_\mu^\dagger(n)\, U_\mu(n) \right\}$$

$$\tilde{S}_\mu(n) = \sum_{\nu \neq \mu} \left( S_{\mu,+\nu}(n) + S_{\mu-\nu}(n) \right)$$

→ $\tilde{S}_\mu(n)$ is no longer a color group element (!), it involves addition of products

→ Consider $N_c = 2$ and drop $\mu$ and $n$ in the notation.

→ We need to sample an $SU(2)$ matrix $U$ from the distribution.

$$dP(U) \sim dU \exp\left\{ \frac{\beta}{2} \operatorname{Re} \operatorname{Tr}\{\tilde{S}^\dagger U\} \right\}$$

where $dU$ is the Haar measure and $\tilde{S}$ is a sum of $SU(2)$ matrices.

→ The crucial property of $SU(2)$ is

that sum of $SU(2)$ matrices is proportional to an $SU(2)$ matrix.
We can normalize:

$$S = \frac{1}{\omega} \tilde{S}, \qquad S \in SU(2)$$

$$\omega^2 = \det \| \tilde{S} \|$$

$$dP(U) \sim dU \exp\left\{ \frac{\beta w}{2} \operatorname{Re} \operatorname{Tr} \{ S^\dagger U \} \right\}$$

→ Change of variables and invariance of the Harr measure:

$$V \equiv S^\dagger U$$

$$dP(V) \sim dV \exp\left\{ \frac{\beta w}{2} \operatorname{Re} \operatorname{Tr} \{ V \} \right\}$$

- Properties of $SU(2)$ Matrix

$$V \equiv a_0 \mathbb{1} + i \vec{a} \cdot \vec{\sigma} = \begin{pmatrix} a_0 + i a_3 & a_2 + i a_1 \\ -a_2 + i a_1 & a_0 - i a_3 \end{pmatrix}$$

$$1 = a_0^2 + |\vec{a}|^2, \quad a_\mu \in [-1,1]$$

$$\mathrm{Re}\,\mathrm{Tr}\{V\} = 2a_0$$

→ Properties of the Measure :

$$dV \sim da_0\, da_1\, da_2\, da_3\, \delta(a_0^2 + |\vec{a}|^2 - 1)$$

$$\sim da_0\, |\vec{a}|^2 d|\vec{a}|\, d(\cos(\theta))\, d\phi\, \delta(|\vec{a}|^2 - (1-a_0^2))$$

$$\sim da_0 \sqrt{1-a_0^2}\, d(\cos\theta)\, d\phi$$

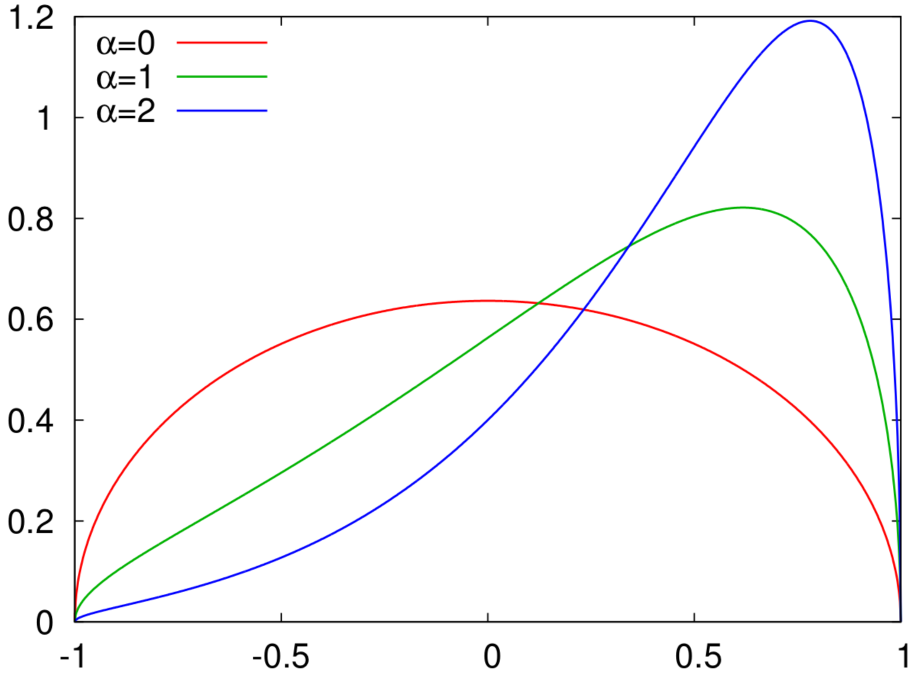
→ The angular part trivially separates :

$$d(\cos\theta)\, d\phi$$

→ We need to sample the following distribution :

$$dP(a_0) \sim da_0 \sqrt{1-a_0^2}\, e^{\alpha a_0}$$

where $\alpha \equiv \beta w$ is a parameter.




★ Fig 1: Distributions for $\alpha = 0, 1, 2$

- PDF & CDF

→ Probability Density Function (PDF)

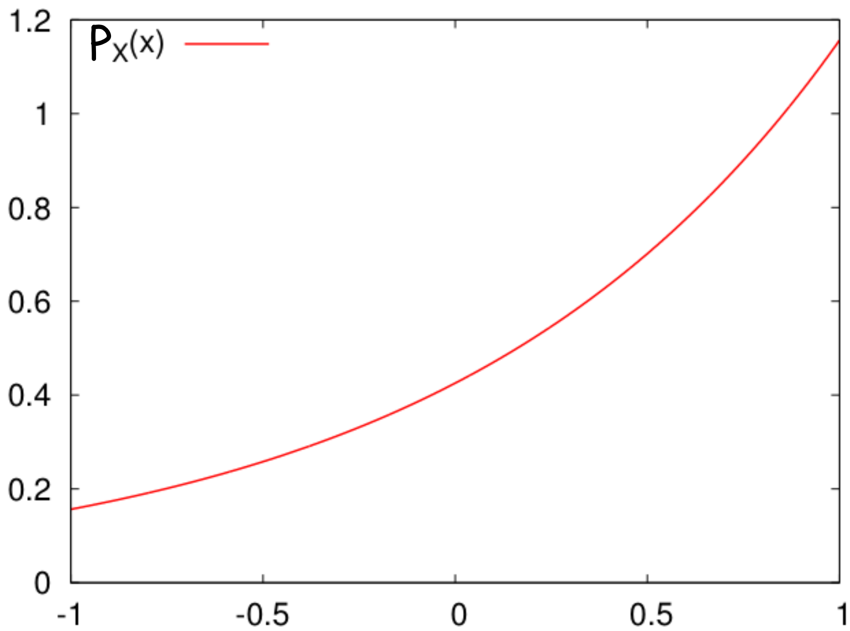
$$p_X(x) = \frac{2\alpha}{\sinh \alpha} e^{\alpha x} \qquad x \in [-1, 1]$$

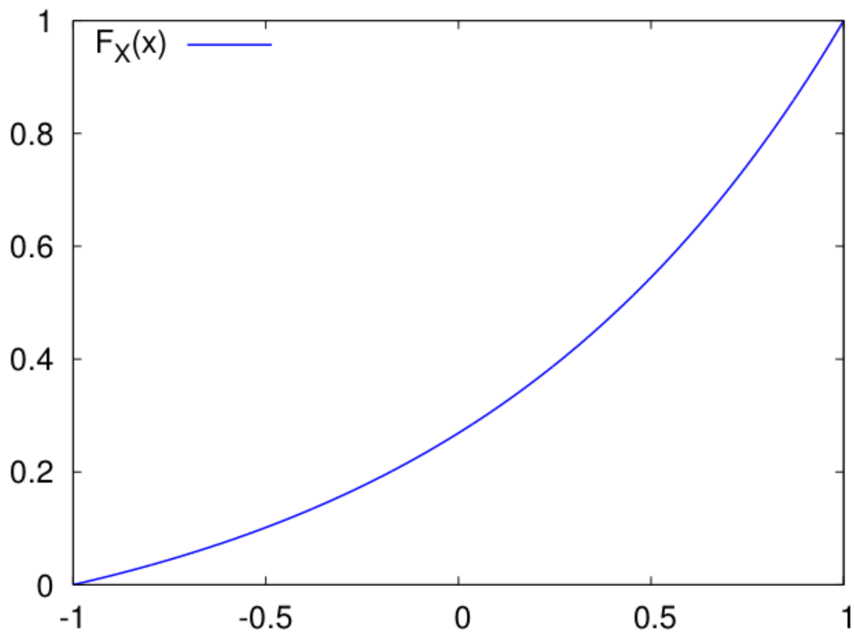
→ cumulative Distribution Function (CDF)

$$F_X(x) \equiv P(X \leq x)$$

$$= \frac{2\alpha}{\sinh\alpha} \int_{-1}^{x} dx' e^{\alpha x'}$$

$$= \frac{2}{\sinh\alpha} \left( e^{\alpha x} - e^{-\alpha} \right)$$




☆ Plots of PDF and CDF

- Inverse CDF Filter

→ Let $X$ be a random variable distributed according to $P_X(x)$.

→ Then $Y \equiv F_X(X)$ is also a random variable.

How is $Y$ distributed?

→ As seen in last class, $Y$ is uniformly distributed in $[0,1]$

→ Then, $X = F_X^{-1}(Y)$ i.e.

The inverse of the CDF acts as a filter that converts a uniform distributed random number into a random variable from the desired distribution $F_X(x)$ eg.

$$X = \frac{1}{\alpha} \ln\left( e^{-\alpha} + \frac{1}{2} \sinh(\alpha Y) \right)$$

→ This can be achieved for a limited no: of P.D.Fs.

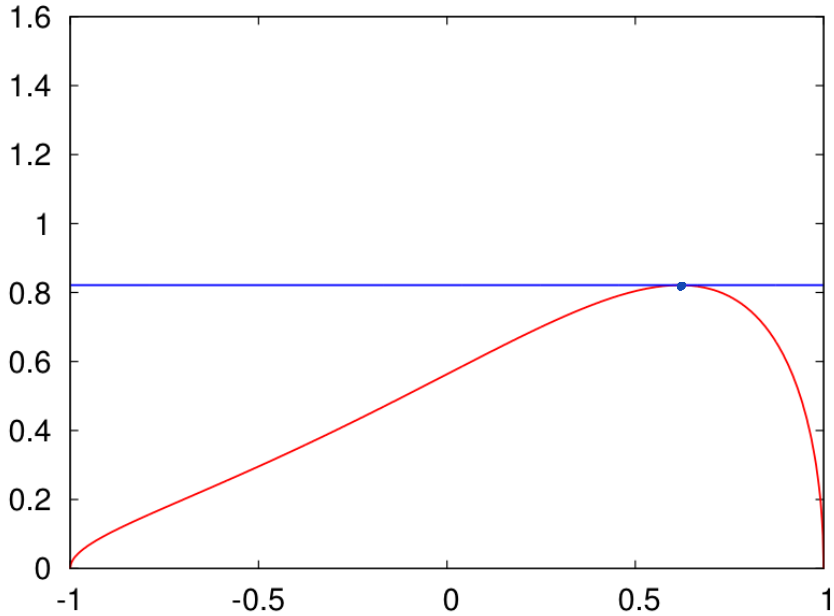
→ In realistic cases, one needs to draw X from a "wrong" distribution $\phi_X(x)$ and then correct by accepting the value with probability :-

$$\sim \frac{P_X(x)}{\phi_X(x)} \quad \text{(The Rejection / Heatbath method)}$$

★ Plot of the $P(a_0)$ Distribution :

→ Consider $dP(a_0) \sim da_0 \sqrt{1-a_0^2}\, e^{\alpha a_0}, \alpha = 1$

→ simplest way :- Draw $a_0$ from $\phi(a_0) \sim 1$, correct with $\sqrt{1-a_0^2}\, e^{\alpha a_0}$ (61% Efficiency)

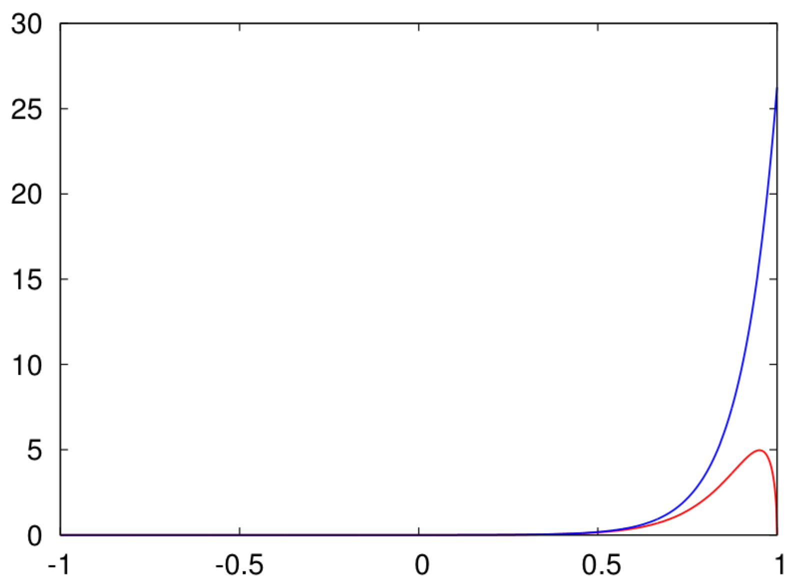


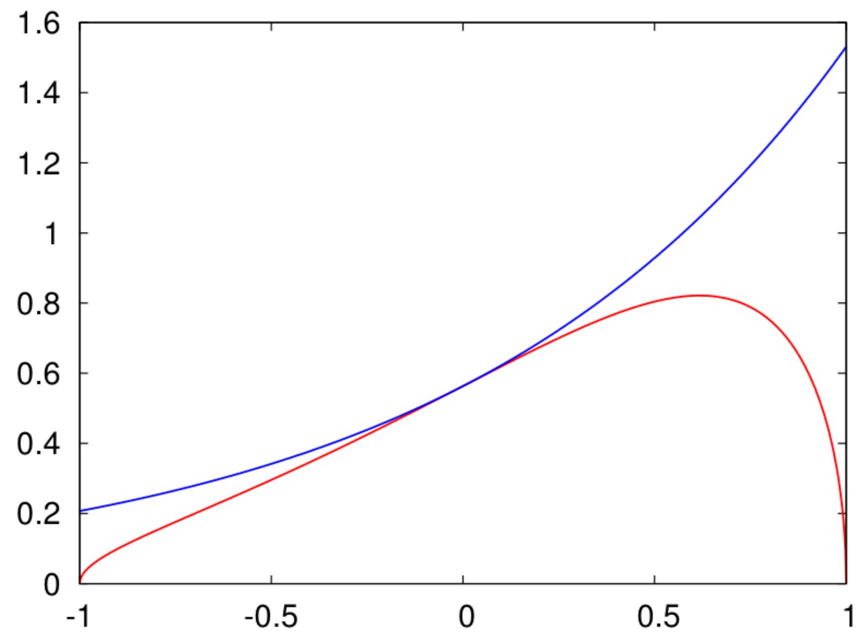

Intersects at one point. (tangential)

★ Plot of $P(x)$, $\phi(x)$ - uniform

→ consider $dP(a_0) \sim da_0 \sqrt{1-a_0^2}\, e^{\alpha a_0}$, $\alpha = 1$

→ Now, Creutz (1980):
Draw $a_0$ from $\phi(a_0) \sim e^{\alpha a_0}$, correct with $\sqrt{1-a_0^2}$ (76% Efficiency)






★ Plot of $P(x)$, $\phi(x) - e^{\alpha a_0}$, $\alpha = 10$

→ Consider $dP(a_0) \sim da_0 \sqrt{1-a_0^2}\, e^{\alpha a_0}$, $\alpha = 10$

→ Draw $a_0$ from $\phi(a_0) \sim e^{\alpha a_0}$, correct with $\sqrt{1-a_0^2}$ (38% efficiency)

→ Other improvements in the $P(a_0)$ Distribution.

→ Further improvements :

- Fabricius, Haan (1984)
- Kennedy, Pendleton (1985)

→ change of variables :

$$r = \sqrt{1-a_0}\ , \quad r \in [0, \sqrt{2}]$$

$$dP(r) \sim dr \sqrt{1 - \frac{1}{2} r^2 r^2 e^{-\alpha r^2}}$$

$$d\varphi(r) \sim dr\, r^2 e^{-\alpha r^2}$$

→ $\varphi(r)$ can be considered as a three-dimensional gaussian distribution with a trivial angular dependence.

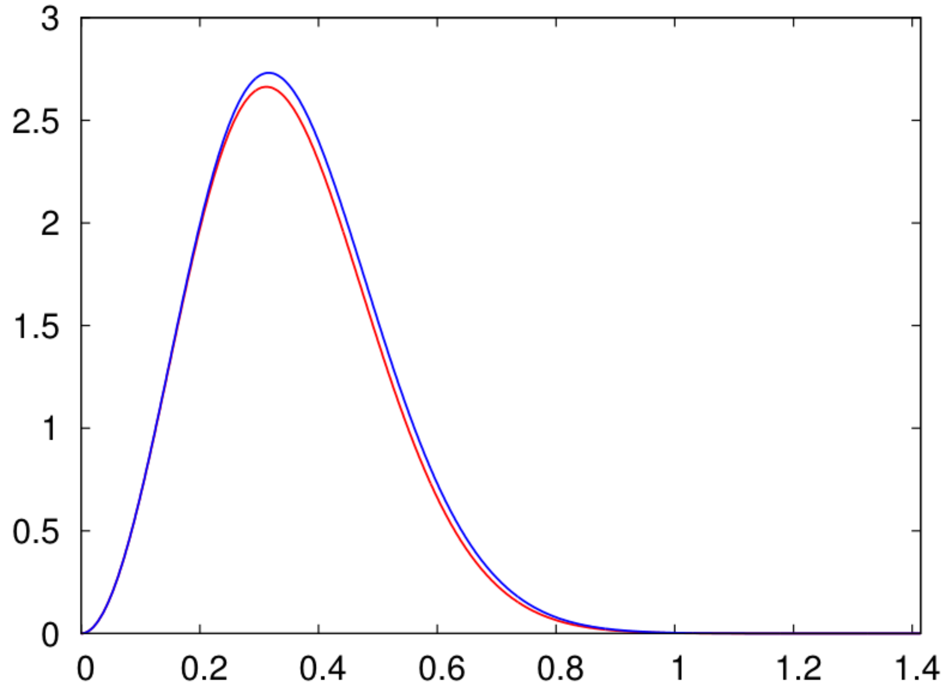
★ Plot of $P(r)$ & $\varphi(r)$ (Improved)

→ consider $P(r) \sim \sqrt{1 - \frac{1}{2} r^2 r^2 e^{-\alpha r^2}}$,

$\alpha = 10$.

→ FHKP : Draw 'r' from $\varphi(r) \sim r^2 e^{-\alpha r^2}$, correct with $\sqrt{1 - r^2/2}$

(96% efficiency)


3
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4

# ★ SU(2) Heatbath Algorithm

① → calculate the sum of the staples $\tilde{S}$ and project to SU(2)

$$S = \tilde{S} / \sqrt{\det \| \tilde{S} \|}$$

② → change of variables $V \equiv S^{+} U$

③ → sample $a_0$ from:

$$dP(a_0) \sim da_0 \sqrt{1 - a_0^2}\, e^{\alpha a_0}$$ (with the Creutz / Improved FHKP Algorithm)

$$\alpha = \beta \sqrt{\det \| \tilde{S} \|}$$

④ → Sample $\theta$ and $\phi$ on a sphere.

⑤ → construct $a_1$, $a_2$, and $a_3$

⑥ → $U = SV$ is a new value for the

link variable.

## ★ $SU(N_c)$ Heatbath Algorithm

Q) How can all this be carried over to arbitrary $N_c > 2$?

→ Updates in $SU(N_c)$ can be built as a sequence of $SU(2)$ subgroups updates, Cabibbo, Marinari (1982)

$$U' = A_m A_{m-1} \cdots A_1 U$$

where,

$U$ is the current $SU(N_c)$ matrix and $A_K$ are $SU(2)$ matrices (embedded in $SU(N_c)$)

→ The staple calculation is carried out in the same way as for the $SU(2)$ case, then for generating the

$A_K$ SU(2) matrix one needs only the part of the total sample matrix that has the same block structure as $A_K$.

→ SU($N_c$) overrelaxation algorithm is similarly built by updating SU(2) subgroups.

→ Multi Hit Metropolis Algorithm, used for Wilson's Gauge action

Metropolis Algorithm w/ single link variable $U_\mu(n)'$. This link is shared by 6 plaquettes and only these 6 plaquettes are affected when changing $U_\mu(n) \rightarrow U_\mu(n)'$.

## Exercises :-

Given a complex scalar field :

$$\varphi(x) = \rho(x)\, e^{i\theta(x)}$$

in the limit $\lambda \to \infty$, where magnitude is frozen $\rho(x) = 1$

The degree of freedom is :

$\theta(x) \in [-\pi, \pi]$. The Action is :-

$$S[\theta] = -2K \sum_{z} \sum_{\hat{\mu}} \mathrm{Re}\left\{ \exp\left(-i\theta_{z+\hat{\mu}}\right) \exp\left(i\theta_{z}\right) \right\}$$

$$= -2K \sum_{z} \sum_{\hat{\mu}} \cos\left(\theta_{z} - \theta_{z+\hat{\mu}}\right)$$

The Partition function is :-

$$Z = \int_{-\pi}^{\pi} \prod_{x} d\theta_x \, e^{-S[\theta]}$$

Goal :- 2D lattice, $V =$ no. of sites.

Build a Markov Chain Monte Carlo process the sample the field configurations (i.e $\theta_x$) & calculate the observable :-

$$L_\varphi \equiv \frac{1}{2V} \sum_{\langle xy \rangle} \exp(-i\theta_x) \exp(i\theta_y)$$

$$\equiv \frac{1}{V} \sum_{z} \sum_{\mu} \mathrm{Re}\left\{ \exp(-i\theta_{z+\hat{\mu}}) \exp(i\theta_z) \right\}$$

$$\equiv \frac{1}{V} \sum_{z} \sum_{\mu} \cos(\theta_z - \theta_{z+\hat{\mu}})$$

# Explaining the setup of the problem :-

Through all the angular,

$$\varphi_z = \rho_z \alpha_z = \rho_z e^{i\theta_z}$$

Complex no.

Radial part (Magnitude)

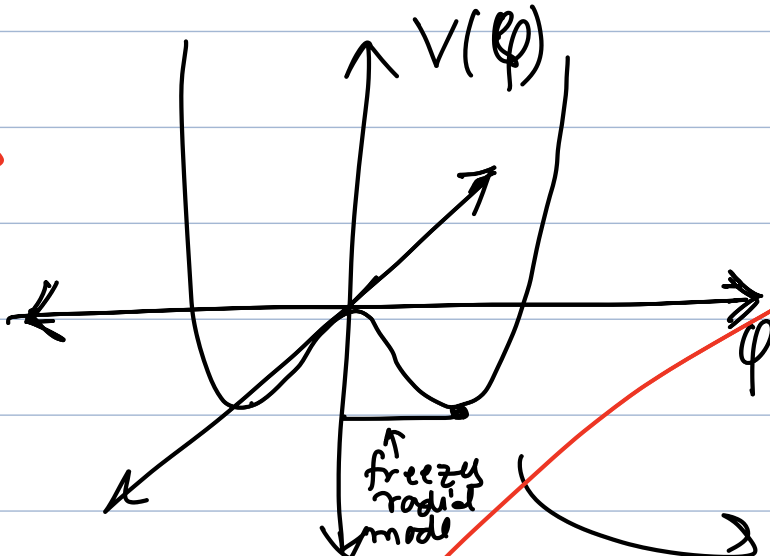
Angular Part

$$|\varphi_z|^2 = \rho_z^2$$

If $\lambda \to \infty$

lattice dimensionless coupling

$$|\varphi_z|^2 = 1 \Rightarrow \rho_z = 1$$




Higg's Mechanism

$\Downarrow$ Simplest

Complex scalar

fipped mass term.

Looking at the action for complex scalar $\phi^4$ theory :-

In limit $\lambda \to \infty$, the interaction term is multiplied by $\infty$, so time deviation of the magnitude from 1 has very large penalty in energy.

→ configurations where the magnitude is not one will be highly suppressed, as they won't contribute to the path integral.
$e^{-S[\varphi]}$

→ This is the limit that freezes the radial mode & we are left with the angular part.

$$S[\varphi] = -2K \sum_z \sum_\mu \cos(\theta_z - \theta_{z+\hat{\mu}})$$

→ we have a degree of freedom per site, which is essentially an angle → (extension of the Ising model)