8 Brief overview of numerical algorithms

see textbooks or Lüscher 1002.4232

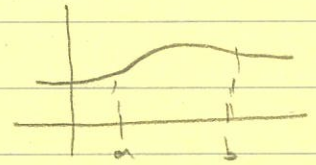
8.1 Importance Sampling

- Lattice regulated functional integral is a regular integral but of very high dimension

$$dim = 2 \times 4 \times 8 \times L^3 \times T \gtrsim 10^{10} \text{ for state-of-the-art}$$

- Monte-Carlo Integration: probabalistic evaluation of integral

- Example: $I = \int_a^b dx\, f(x)$



estimate as

$$\bar{I} = \frac{(b-a)}{N} \sum_{i=1}^{N} f(x_i) = (b-a)\langle f(x)\rangle$$

where x_i are uniformly distributed random points in $[a,b]$

Will converge as $N \to \infty$

$$I = \bar{I} + \frac{\sigma}{\sqrt{N}} \qquad \sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(f(x_i) - \langle f\rangle\right)^2$$

- This will work well for slowly varying, smooth function
- Not so well if function varies sharply in integration region

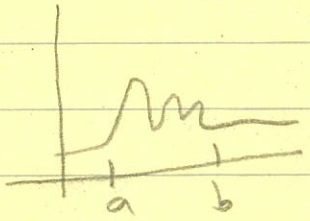
- Importance sampling: change of variables to make integrand slowly varying

Let $f(x)$ very rapidly over $[a,b]$

Try to find $f(x) = g(x)p(x)$ where

$p(x)$ positive and bounded

$g(x)$ slowly varying on $[a,b]$



$$I = \int_a^b f(x)\,dx = \int_{y_a}^{y_b} dy\, g(y) \to \bar{I} = \frac{1}{N} \sum_{i=1}^{N} g(y_i)$$

y_i uniformly distributed on $[y_a, y_b]$

where

$$p(x)\,dx = dP(x) = dy \qquad y = P(x) \text{ change of variables}$$

$$\bar{I} = \frac{1}{N} \sum_{i=1}^{N} g(x_i)\, V_{ab}$$

with x_i random numbers in $[a,b]$ distributed according to probability density $\frac{1}{V_{ab}} p(x)\,dx$

with $V_{ab} = \int_a^b p(x)\,dx$

$$\int_0^{\pi} \sin\frac{\pi x}{10}\,dx \quad , \quad \int_0^{10} e^{-x^2}\,dx$$

- Try some examples!

- In QFT we want to evaluate partition function and correlation functions

$$Z = \int \mathcal{D}\varphi\, e^{-S[\varphi]}$$

$$C_n = \langle \phi_1(x_1) \dots \phi_n(x_n) \rangle = \frac{1}{Z} \int \mathcal{D}\varphi\, \phi_1 \dots \phi_n\, e^{-S[\varphi]}$$

Very sharply peaked at small action.

⇒ Importance sample with probability density

$$\frac{1}{Z} e^{-S[\varphi]} \mathcal{D}\varphi$$

(works since $e^{-S[\varphi]}$ positive and bounded) (ignoring μ_B, Θ term)

but how do we sample according to a distribution

→ Evaluate correlation function

$$C_n = \frac{1}{N} \sum_{i=1}^{N} \phi_1^{[i]} \dots \phi_n^{[i]} + \Theta\left(\frac{1}{\sqrt{N}}\right) \quad \checkmark$$

where $\phi^{[i]}$ corresponds to the i^{th} sample

→ How do we produce a set of samples with the desired probability distribution?

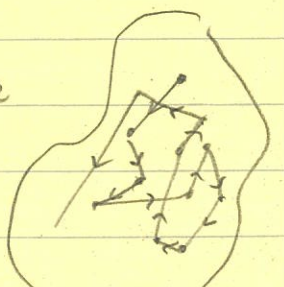
⇒ Markov chain : a stochastic sequence of configurations of the d.o.f. of a system defined by a transition probability $T(C, C')$ from 1 state to another

states C, C' $\quad C \xrightarrow{T(C,C')} C'$

transition only depends on initial state and final state

Properties 1) $\sum_{C'} T(C', C) = 1$

2) $0 \le T(C, C') \le 1 \quad \forall C', C$



NB: each state corresponds to values for d.o.f. @ all sites in lattice

- Desire that particular MC has the target probability distribution as a fixed point (ie once we have the desired prob dist, further application does not destroy it

$$\sum_C P_{eq}(C)\, T(C',C) = P_{eq}(C') \qquad (*)$$

- Ergodic : possible to get to every possible state
- MC with these properties is guaranteed to arrive at $P_{eq}(C)$ after many steps starting from an arbitrary state after which new states are produced in sequence following the desired prob distribution

Define a distance between distribution through an ensemble of states

$$\|E' - E\| = \sum_C |P'(C) - P(C)|$$

then $E' = \{ C = T(C,C')P(C') \;\; \forall C' \in E \}$

$$\|E' - E_{eq}\| = \sum_C |P'(C) - P_{eq}(C)|$$

$$= \sum_C \sum_{C'} |T(C,C')[P(C') - P_{eq}(C')]| \quad \text{use } (*)$$

$$\leq \sum_{CC'} T(C,C')\,|P(C') - P_{eq}(C')| \quad T > 0$$

$$= \sum_{C'} |P(C') - P_{eq}(C')| \quad \sum_C T(C,C') = 1$$

$$= \|E - E_{eq}\|$$

ie ensemble E' is closer to E_{eq} than E was

- In order to have a fixed point it is sufficient (but not necessary) to require <u>detailed balance</u>

$$T(C',C)\,P_{eq}(C) = T(C,C')\,P_{eq}(C')$$

Since $\sum_{C'} T(C,C') = 1$, $(*)$ implies the balance condition must hold

$$\sum_C T(C',C)\,P_{eq}(C) = \sum_C T(C,C')\,P_{eq}(C')$$

- Need to find an algorithm to implement $T(C,C')$ with these properties that reaches the equilibrium distribution quickly

Metropolis algorithm: algorithm satisfying detailed balance that generically works

Step 1: choose a updated state C' according to a given transition prob $T_0(C', C)$ that can be implemented efficiently

Step 2: accept or reject state C' according to the acceptance prob

$$T_{Ac}(C', C) = \min\left(1, \frac{T_0(C, C')\, P_{eq}(C')}{T_0(C', C)\, P_{eq}(C)}\right)$$

$T = T_0 T_{Ac}$ satisfies detailed balance

$$\begin{aligned} T(C', C)\, P_{eq}(C) &= T_0(C', C) \min\left(1, \frac{T_0(C, C')\, P_{eq}(C')}{T_0(C', C)\, P_{eq}(C)}\right) P_{eq}(C) \\ &= \min\left(T_0(C', C) P_{eq}(C),\ T_0(C, C') P_{eq}(C')\right) \\ &= T_0(C, C') \min\left(\frac{T_0(C', C) P_{eq}(C)}{T_0(C, C') P_{eq}(C')}, 1\right) P_{eq}(C') \\ &= T(C, C') P_{eq}(C') \end{aligned}$$

→ Common choice is to have $T_0(C, C') = T_0(C', C)$ symmetric

$$\Rightarrow T_{Ac}(C', C) = \min\left(1, P_{eq}(C')/P_{eq}(C)\right)$$

or since our target prob dist is $\frac{1}{Z} e^{-S}$

$$T_{Ac}(C', C) = \min\left(1, e^{-\Delta S}\right) \qquad \Delta S = S(C') - S(C)$$

change in action

→ Metropolis algorithm efficient for ultralocal actions (pure gauge theory, spin systems etc)

→ Not effective for theories with fermions since after integrating over fermions effective action that results looks non-local

→ For QCD calculations the state-of-the-art algorithm is Hybrid Monte Carlo (HMC) - beyond current scope/time

→ For gauge theory, update T_0 can be multiply a single gauge link by a random group element $U_\mu'(n) = X U_\mu(n)$ X random SU(N) matrix (close to $\mathbb{I}$), repeat steps $\forall$ links

8.2 Measurements

- We now have an algorithm to generate an ensemble of states representative of the correct probability distribution. Measurement of correlation functions is then simply an average of the field variables that define the correlator over the ensemble of states
- In QCD, eg average value of plaquette (value of link in ith state)

$$\langle \square \rangle = \frac{1}{N} \sum_{i=1}^{N} \sum_{n} \sum_{\mu<\nu} \mathrm{tr}\left[U_\mu^{[i]}(n)\, U_\nu^{[i]}(n+\hat{\mu})\, U_\mu^{+[i]}(n+\hat{\nu})\, U_\nu^{+[i]}(n)\right]$$

or Wilson loop (large closed path of links)

- For fermionic observables, we have ($M[U]$)

$$\langle \bar{\psi}_{x_1}\psi_{x_1}\bar{\psi}_{x_2}\psi_{x_2} \ldots \rangle = \int \mathcal{D}U \mathcal{D}\bar{\psi} \mathcal{D}\psi\ \bar{\psi}_{x_1}\psi_{x_1}\bar{\psi}_{x_2}\psi_{x_2}\, e^{-S_{\text{gauge}} + \bar{\psi} M \psi}$$

$$\int dU \left(\frac{1}{M}\right)_{x_1x_2}\left(\frac{1}{M}\right)_{x_2x_1} + \left(\frac{1}{M}\right)_{x_1x_1}\left(\frac{1}{M}\right)_{x_2x_2} + \ldots$$

$$\to \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{M[U^{[i]}]}\right)_{x_1x_2}\left(\frac{1}{M[U^{[i]}]}\right)_{x_2x_1} + \ldots$$

→ Need inverse of matrix $M[U^{[i]}]$ for the ith state

$\dim[M] = (12V)\times(12V) \sim 10^{12}\times 10^{12}$!!

but very sparse

→ Recast problem as linear system, solve

$Mx = b$ for $b = \begin{pmatrix}1\\0\\0\\\vdots\end{pmatrix}, \begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix}, \begin{pmatrix}0\\0\\1\\\vdots\end{pmatrix} \ldots$

$\Rightarrow x = M^{-1}b$ gives row of M^{-1}

→ Best direct solver methods (LU factorisation, ...) take $\mathcal{O}((12V)^3)$ op

→ Iterative methods which construct a sequence of improving approximation to x have a cost $\mathcal{O}(12V)$ (exact result)

$x_0 \to x_1 \to x_2 \to \ldots x_n$ with $\|x_{i+1} - \hat{x}\| < \|x_i - \hat{x}\|$

→ Many specific algorithms of generic type that are called Krylov space methods

- One example is <u>conjugate gradient algorithm</u> (CG) which works for Hermitian systems

[M is not Hermitian, but $M^{-1} = M^{\dagger}(MM^{\dagger})^{-1}$ so we can work with $A = M^{\dagger}M$ which is Hermitian

→ Proceeds via minimizing a quadratic form

$$f(\vec{x}) = \frac{1}{2}\vec{x}^{T} A \vec{x} - \vec{x}^{T}\vec{b}$$

$$\vec{\nabla} f = 0 \quad @ \quad A\vec{x}^{*} - \vec{b} = 0$$

by iterating the minimization in orthogonal direction

find α_i such that $f(\vec{x}^{(i)} + \alpha_i \vec{p}^{(i)})$ is minimized

with $\vec{x}^{(i+1)} = \vec{x}^{(i)} + \alpha_i \vec{p}^{(i)}$

and $\vec{p}^{(i)}$ orthogonal to all previous $\vec{p}^{(j)}$ in the A-norm

$$\vec{p}^{(i)} A \vec{p}^{(j)} \sim \delta_{ij}$$

The space $K^{(i)} = \mathrm{span}\{\vec{p}^{(0)} = \vec{b} - A\vec{x}^{(0)}, \vec{p}^{(1)}, \vec{p}^{(2)}, \ldots, \vec{p}^{(i)}\}$ is the i^{th} Krylov subspace

→ The sequence $\vec{x}^{(0)} \to \vec{x}^{(1)} \to \ldots \to \vec{x}^{(N)} = x^{*}$

arrives at the exact solution after $N = \dim A$ steps

but typically $|\vec{x}^{(i)} - \vec{x}^{*}| <$ machine precision much earlier

[<u>CG Algorithm</u> Guess $\vec{x}_0$

$\vec{r}_0 = \vec{b} - A\vec{x}_0$

$\vec{p}_0 = \vec{r}_0$

do $k = 0, \ldots, N$

$\alpha_k = \vec{r}_k^{\dagger} \cdot \vec{r}_k \,/\, \vec{p}_k^{\dagger} A \vec{p}_k$

$\vec{x}_{k+1} = \vec{x}_k + \alpha_k \vec{p}_k$

$\vec{r}_{k+1} = \vec{r}_k - \alpha_k A \vec{p}_k$

if $\|\vec{r}_{k+1}\| < tol\, \|\vec{b}\|$ stop & return $\vec{x}_{k+1}$

$b_k = \vec{r}_{k+1}^{\dagger} \cdot \vec{r}_{k+1} \,/\, \vec{r}_k^{\dagger} \cdot \vec{r}_k$

$\vec{p}_{k+1} = \vec{r}_{k+1} + b_k \vec{p}_k$

→ All implemented as simple linear algebra operations on large vectors

→ Typically the number of iterations required for convergence to a particular precision is set by the condition # of the matrix A

$$C(A) = \frac{\lambda_{max}}{\lambda_{min}}$$ ratio of max and min eigenvalues

Since in our case $A = M^\dagger M = (D^2 - m_q^2)$, λ_{min} is set by quark mass m_q (λ_{max} is set by cutoff) and approaching light quark masses makes the numerical problem harder & harder

→ Many other methods that improve on CG also exist, some of which scale much better with m_q

→ This linear algebra problem is a very large part of the cost of LQCD calculations and around the world many 10^9 cpu hours are being spent on it!!

5 Application 1 : Hadron Spectroscopy

- Given the ability to address QCD at low energies, the most basic quantities to calculate are the masses of the stable hadrons of QCD
To do this, we construct two representations of hadronic correlation functions
- Let $\mathcal{O}_H[\bar{\psi},\psi,U]$ be an operator with the quantum numbers of the hadron of interest.
Eg $\mathcal{O}_{\pi^+} = \bar{u}(x)\gamma_5 d(x)$
- We can construct

$$C_H(\vec{x},t) = \langle 0|\mathcal{O}_H(\vec{x},t)\,\mathcal{O}_H^\dagger(0,0)|0\rangle = \frac{1}{Z}\mathrm{Tr}\left[e^{-\hat{H}(T-t)}\hat{\mathcal{O}}_H(\vec{x})\,e^{-\hat{H}t}\hat{\mathcal{O}}_H^\dagger(\vec{0})\right]$$
$$= \frac{1}{Z}\sum_\rho \langle\rho|e^{-\hat{H}(T-t)}\hat{\mathcal{O}}_H(x)e^{-\hat{H}t}\hat{\mathcal{O}}_H^\dagger(\vec{0})|\rho\rangle$$

(a sum over all states in Hilbert space). Now, using completeness

$$= \frac{1}{Z}\sum_\rho\sum_\sigma \langle\rho|e^{-\hat{H}(T-t)}\hat{\mathcal{O}}_H(\vec{x})|\sigma\rangle\langle\sigma|e^{-\hat{H}t}\mathcal{O}_H^\dagger(\vec{0})|\rho\rangle$$
$$= \frac{1}{Z}\sum_{\rho,\sigma} e^{-E_\rho(T-t)}e^{-E_\sigma t}\langle\rho|\hat{\mathcal{O}}_H(\vec{x})|\sigma\rangle\langle\sigma|\mathcal{O}_H^\dagger(\vec{0})|\rho\rangle$$

Similarly $Z = \sum_\rho \langle\rho|e^{-\hat{H}T}|\rho\rangle = \sum_\rho e^{-E_\rho T} = e^{-E_0 T}\left(1+e^{-\Delta E_1 T}+e^{-\Delta E_2 T}+\ldots\right)$
where $\Delta E_n = E_n - E_0$ with E_0 the vacuum energy

$$\Rightarrow\quad C_H(\vec{x},t) = \frac{\sum_{\rho,\sigma}\langle\rho|\mathcal{O}_H|\sigma\rangle\langle\sigma|\mathcal{O}_H^\dagger|\rho\rangle\, e^{-(T-t)\Delta E_\rho}e^{-\Delta E_\sigma t}}{1+e^{-\Delta E_1 T}+e^{-\Delta E_2 T}+\ldots}$$

Now, taking T large, only $|\rho\rangle = |0\rangle$ (lowest state) contributes

$$C_H(\vec{x},t) \xrightarrow[T\to\infty]{} \sum_\sigma \langle 0|\mathcal{O}_H|\sigma\rangle\langle\sigma|\mathcal{O}_H^\dagger|0\rangle e^{-\Delta E_\sigma t} \qquad (\dagger)$$

$\Rightarrow$ Using the functional integral form for the correlator

$$C_H(\vec{x},t) = \int \mathcal{D}U\mathcal{D}\bar{\psi}\mathcal{D}\psi\, e^{-S_{\text{gauge}}+\int\bar{\psi}M\psi}\,\bar{\psi}_u(x,t)\gamma_5\psi_d(x,t)\,\bar{\psi}_d(\vec{0},0)\gamma_5\psi_u(\vec{0},0)$$
$$= \int \mathcal{D}U \det M_u \det M_d\, e^{-S_{\text{gauge}}+\mathrm{tr}\left[M_u^{-1}(x,0)\gamma_5 M_d^{-1}(0,x)\gamma_5\right]}$$
$$\longrightarrow \frac{1}{N_{cf}}\sum_{c=1}^{N_{cf}} \mathrm{tr}\left[M_u^{-1}[U_c]\gamma_5 M_d^{-1}[U_c]\gamma_5\right] \qquad (\circledast)$$

where there are N_{cf} configs in ensemble $\{U_1, U_2 \ldots U_{N_{cf}}\}$

- Eq (*) tells us what to calculate in terms of quark propagators $M^{-1}[U]$ and (*) tells us how to interpret the result

→ Let us explicitly $\sum_{\vec{x}}$ in both cases. Then

$$C(t) \quad \sum_{\vec{x}} C(\vec{x},t) = \sum_{\sigma(\vec{p}=0)} |Z_\sigma|^2 e^{-M_\sigma t}$$

where the sum is restricted to states of momentum $\vec{p}=0$ and

$$Z_\sigma = \langle 0 | \mathcal{O}_H | \sigma(\vec{p}=0) \rangle$$

→ These correlation functions determine the energies of states for which $Z_\sigma \neq 0$, that is states that can be created from the vacuum by the operator $\hat{\mathcal{O}}_H^\dagger$

→ There are many choices for the <u>interpolating operators</u> $\mathcal{O}_H$ that have the same quantum numbers.

Eg $\hat{\mathcal{O}}_{\pi^+}(x) = \bar{\tilde{u}}(x) \gamma_4 \gamma_5 \tilde{d}(x)$

where $\sim$ denotes that the quark field has been smeared in some way $\tilde{q}(x) = \sum_y f(x,y) q(y)$ for a given $f(x,y)$

→ Taking $t \to \infty$

$$C(t \to \infty) \to |Z_1|^2 \exp(-M_1 t) + \text{small}$$

where M_1 is the lightest state with the right quantum numbers.

→ It is possible (although difficult) to extract excited states by fitting multiple exponentials to the $C(t)$ time series. A more promising method is to use a variational approach in which we choose a set of interpolating operators all of the same quantum numbers

$$\{ \mathcal{O}_1, \mathcal{O}_2, \ldots \mathcal{O}_J \}$$

and use them to construct a matrix of correlators

with elements

$$C_{ij}(t) = \frac{1}{Z} \mathrm{Tr}\left[e^{-\hat{H}(T-t)} \hat{\mathcal{O}}_i e^{-\hat{H}t} \hat{\mathcal{O}}_j^\dagger \right]$$

$$\xrightarrow[T\to\infty]{} \sum_\sigma \langle 0|\mathcal{O}_i|\sigma\rangle\langle\sigma|\mathcal{O}_j^\dagger|0\rangle\, e^{-E_\sigma t} = \sum_\sigma Z_{ij}^{(\sigma)} e^{-E_\sigma t}$$

- Each component has the same set of energies but different overlap factors. Some will overlap well onto the lowest state while others may overlap best onto excited states
- Solving the eigenvalue problem associated with this matrix gives eigenvalues $\{ e^{-E_1 t}, e^{-E_2 t}, \dots \}$ up to exponentially supressed terms

[We solve the generalised eigenvalue problem
$$C^{-1/2}(t_0)\, C(t)\, C^{\dagger -1/2}(t_0)\, v = \lambda(t)\, v$$
as it is numerically more stable]

6 Application 2: Hadron Structure

- More complex correlation functions allow us access to matrix elements in hadronic states
- We will look at the isovector axial coupling of the nucleon defined by the matrix element

$$\langle N(\vec{p})|\, \bar{\psi}\gamma_\mu\gamma_5\tau_3\psi \,|N(\vec{p})\rangle \sim g_A\, \bar{u}_p(\vec{p})\gamma_\mu\gamma_5 u_p(\vec{p})$$

($\bar{\psi}\gamma_\mu\gamma_5\tau_3\psi = J^3_{\mu 5}$)

- To extract this, we consider a 3 point function

$$C_3^\mu(t,\tau) = \frac{1}{Z}\sum_{\vec{x},\vec{y}} \mathrm{Tr}\left[e^{-\hat{H}(T-t)}\, \mathcal{O}_N(\vec{x})\, e^{-\hat{H}(t-\tau)}\, \hat{J}^3_{\mu 5}(\vec{y})\, e^{-\hat{H}\tau}\, \mathcal{O}_N^\dagger(0) \right]$$

$$\xrightarrow[T\to\infty]{} \sum_{\vec{x},\vec{y}} \sum_{n,m} \langle 0|\hat{\mathcal{O}}_N|n\rangle\langle n|e^{-H(t-\tau)}\hat{J}^3_{\mu 5} e^{-\hat{H}\tau}|m\rangle\langle m|\hat{\mathcal{O}}_N^\dagger|0\rangle$$

$$= \sum_{\vec{x},\vec{y}} Z_n Z_m^\dagger e^{-E_n(t-\tau)} e^{-E_m\tau} \langle n|J^3_{\mu 5}|m\rangle$$

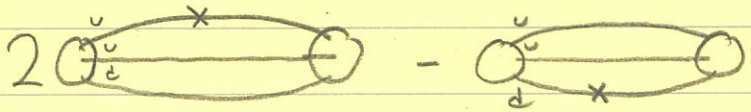
where again, the $Z_n = \langle 0|\hat{\mathcal{O}}_N|n\rangle$ are overlap factors and the sums are over all states with the correct quantum numbers and with $\vec{p}=0$

- The Z_n's and E_n's can be determined from an appropriate two point function

$$C_N(t) = \frac{1}{Z}\sum_x \mathrm{Tr}\left[e^{-H(T-t)}\mathcal{O}_N(\vec{x})\,e^{-Ht}\mathcal{O}_N^\dagger(0)\right]$$
$$\to \sum_n |Z_n|^2 e^{-E_n t}$$

So given measurements of $C_3^\mu(t,\tau)$, the matrix element of the axial current in the nucleon ground state can be determined

- We also need to determine the expression for $C_3^\mu(t,\tau)$, $C_2(t)$ in terms of quark propagators. This is straightforward and best shown diagramatically



$$\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where these skeleton diagrams show the way in which the quark propagators are tied/contracted together

NB: for baryons, the interpolating operators contain a colour antisymmetric Levi-Civita tensor and a Dirac structure

$$\mathcal{O}_N \sim \varepsilon^{abc}\left(u^a C\gamma_5 d^b\right) u^c_\alpha$$

(c: colour index; α: spin index)

7 Summary

- We have covered all the basics needed to get started in lattice QCD
- Go out and use them!