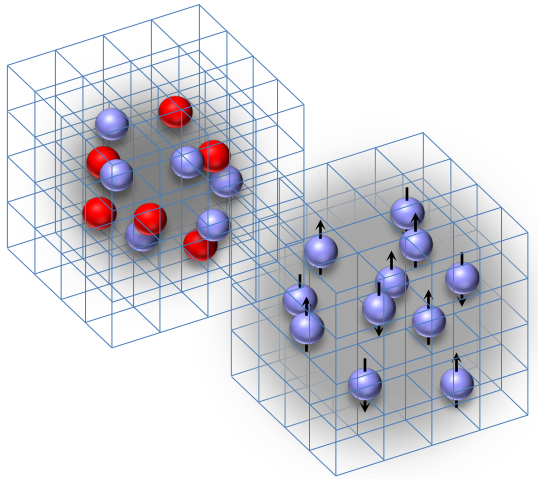
Nuclear Lattice Simulations
Lecture 1: Lattice Formalism and Monte Carlo Methods

Dean Lee (leed@frib.msu.edu)
Facility for Rare Isotope Beams
Michigan State University
Nuclear Lattice EFT Collaboration
June 23, 2020



ACPTP Nuclear Physics School 2020
June 22-26, 2020

Outline

Introduction and history

Grassmann path integrals

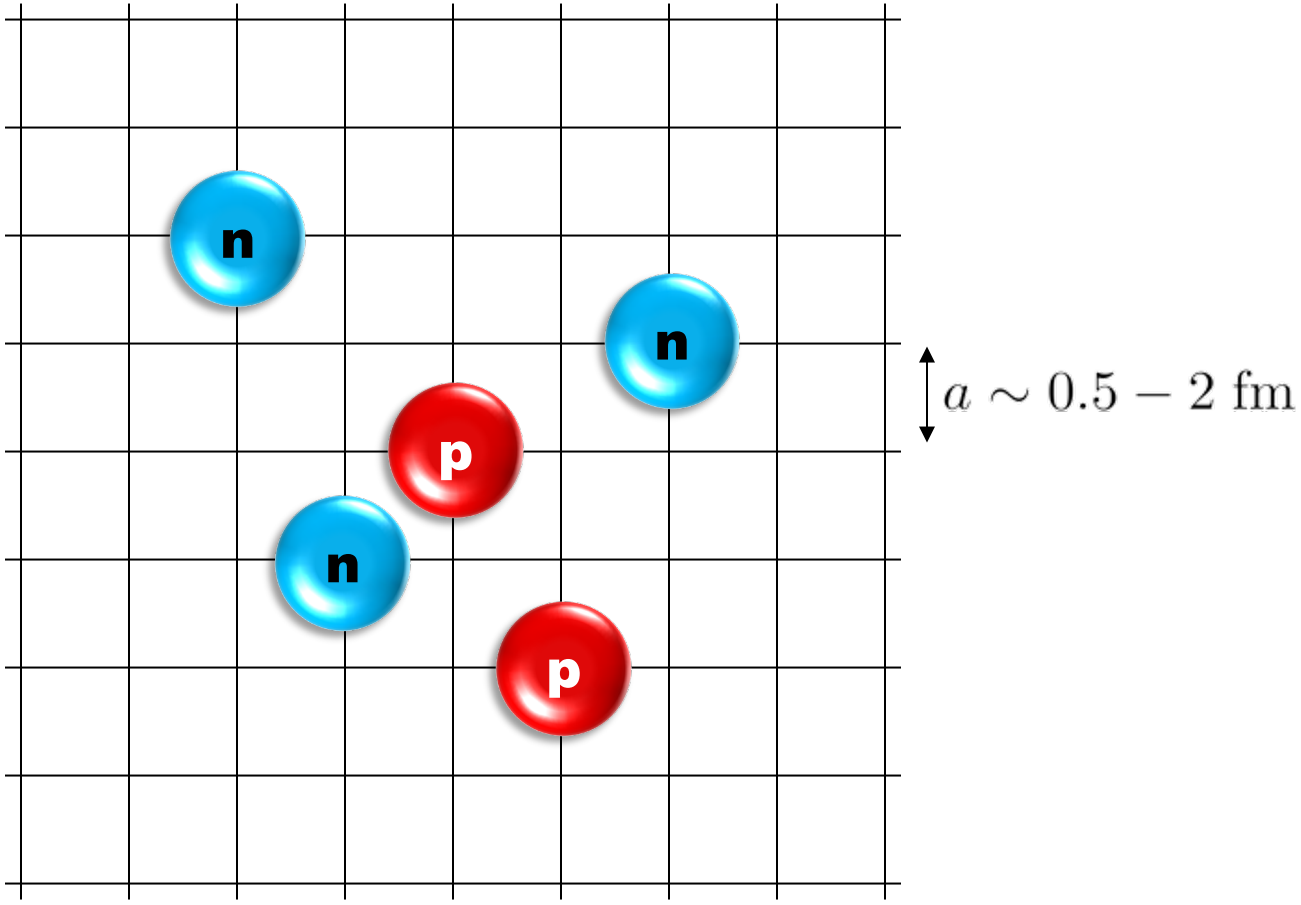
Transfer matrix operators

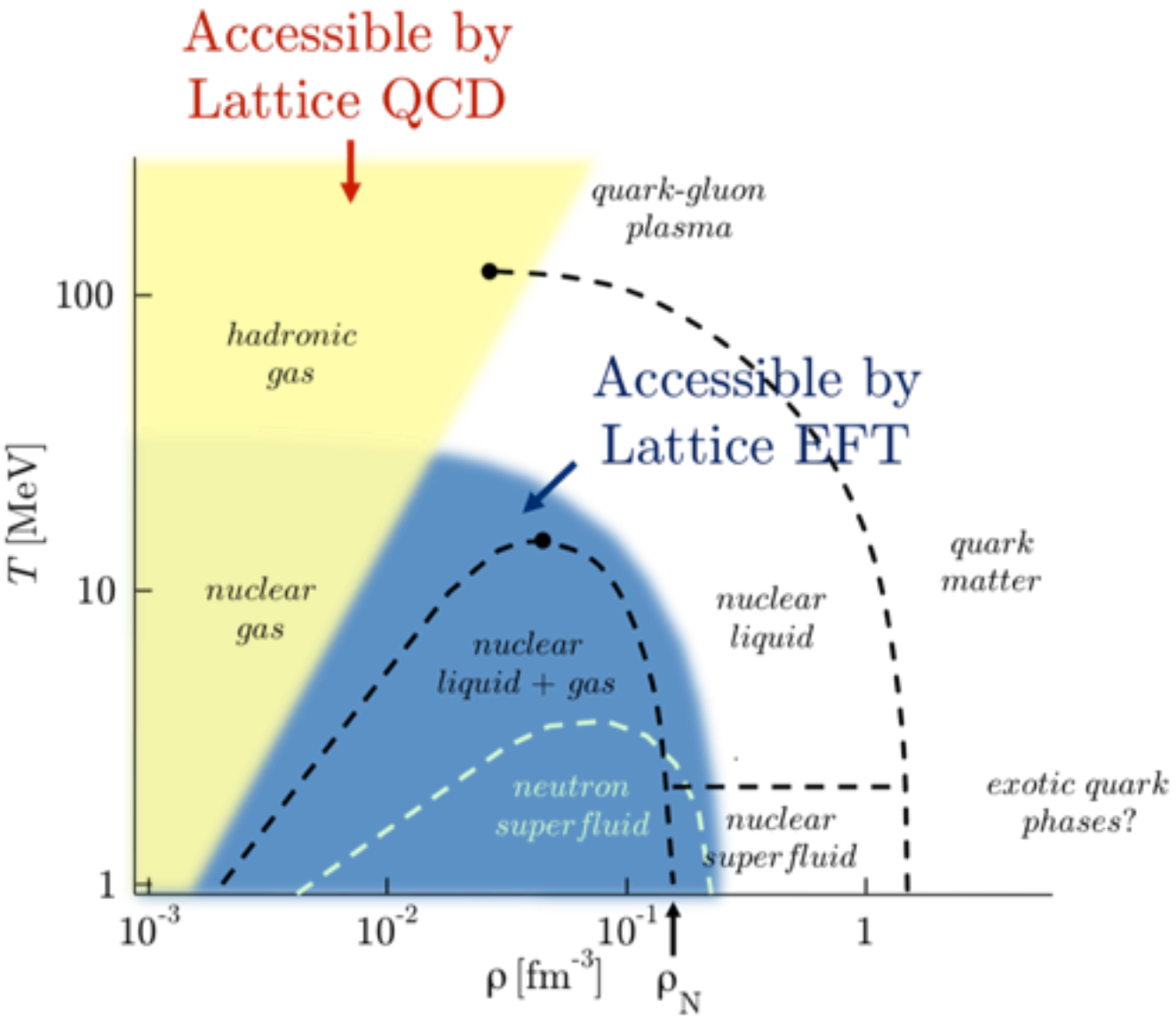
Auxiliary fields

Monte Carlo methods

Markov chain Monte Carlo

Nuclear lattice effective field theory




Accessible by
Lattice QCD
quark-gluon
plasma
hadronic
gas
Accessible by
Lattice EFT
quark
matter
nuclear
gas
nuclear
liquid
nuclear
liquid + gas
neutron
superfluid
nuclear
superfluid
exotic quark
phases?
T [MeV]
100
10
1
10^-3
10^-2
10^-1
1
ρ [fm^-3]
ρ_N

Early lattice EFT papers on nuclear physics

Brockman, Frank, PRL 68 (1992) 1830

Shushpanov, Smilga, Phys. Rev. D59 (1999) 054013

Müller, Koonin, Seki, van Kolck, PRC 61 (2000) 044320

Lewis, Ouimet, PRD 64 (2001) 034005

Chandrasekharan, Pepe, Steffen, Wiese, JHEP 12 (2003) 35

D.L., Borasoy, Schaefer, PRC 70 (2004) 014007

Early lattice EFT papers on cold atoms

Chen, Kaplan, PRL 92 (2004) 257002

Wingate, cond-mat/0502372

D.L., Schaefer, PRC 73 (2006) 015202

Bulgac, Drut, Magierski, PRL 96 (2006) 090404

Burovski, Prokofev, Svistunov, PRL 96 (2006) 160402

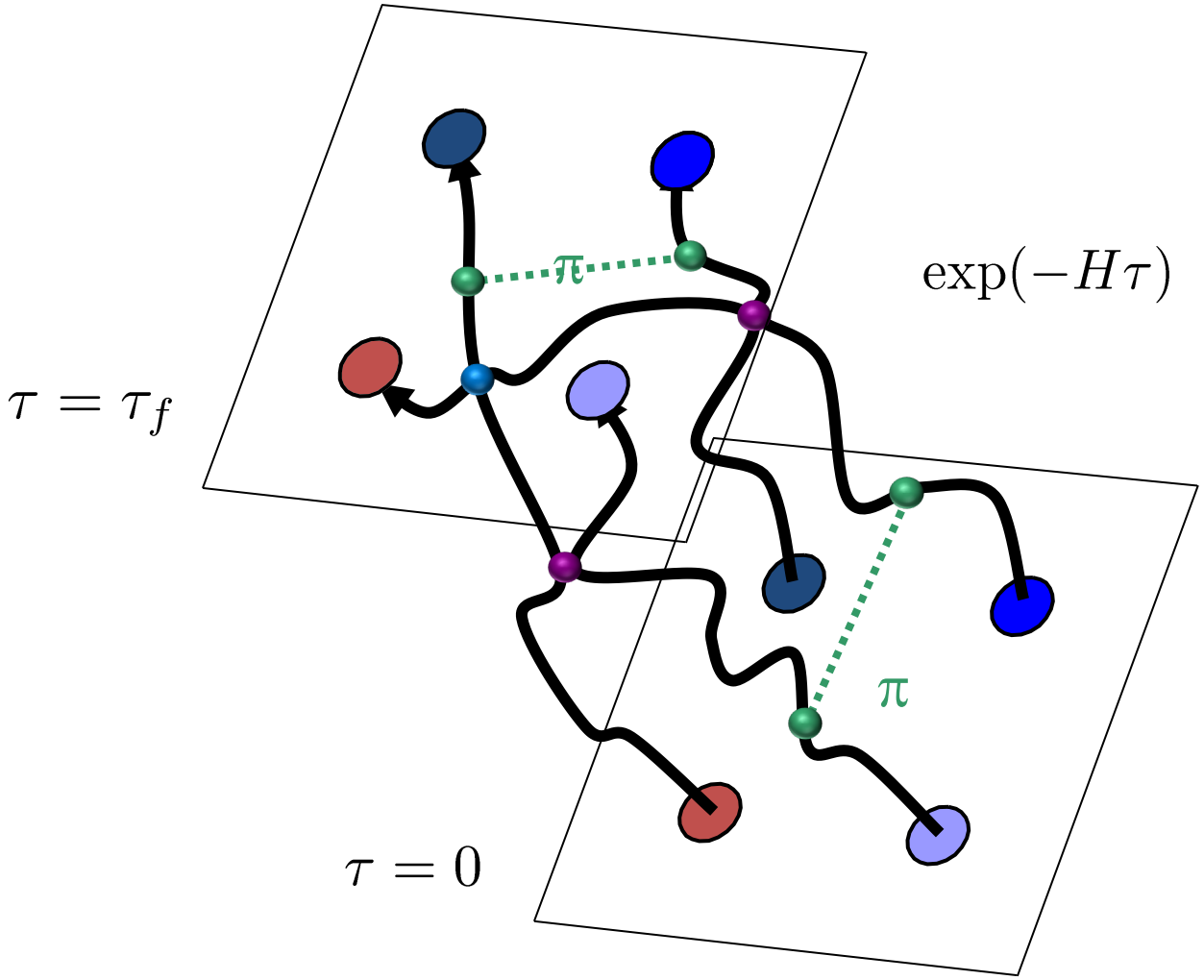
Review articles and textbooks

D. L., Prog. Part. Nucl. Phys. 92 (2009) 117

Drut, Nicholson, J. Phys. G: Nucl. Part. Phys. 40 (2013) 043101

Lähde, Meißner, "Nuclear Lattice Effective Field Theory: An Introduction", Springer (2019)

Euclidean time projection



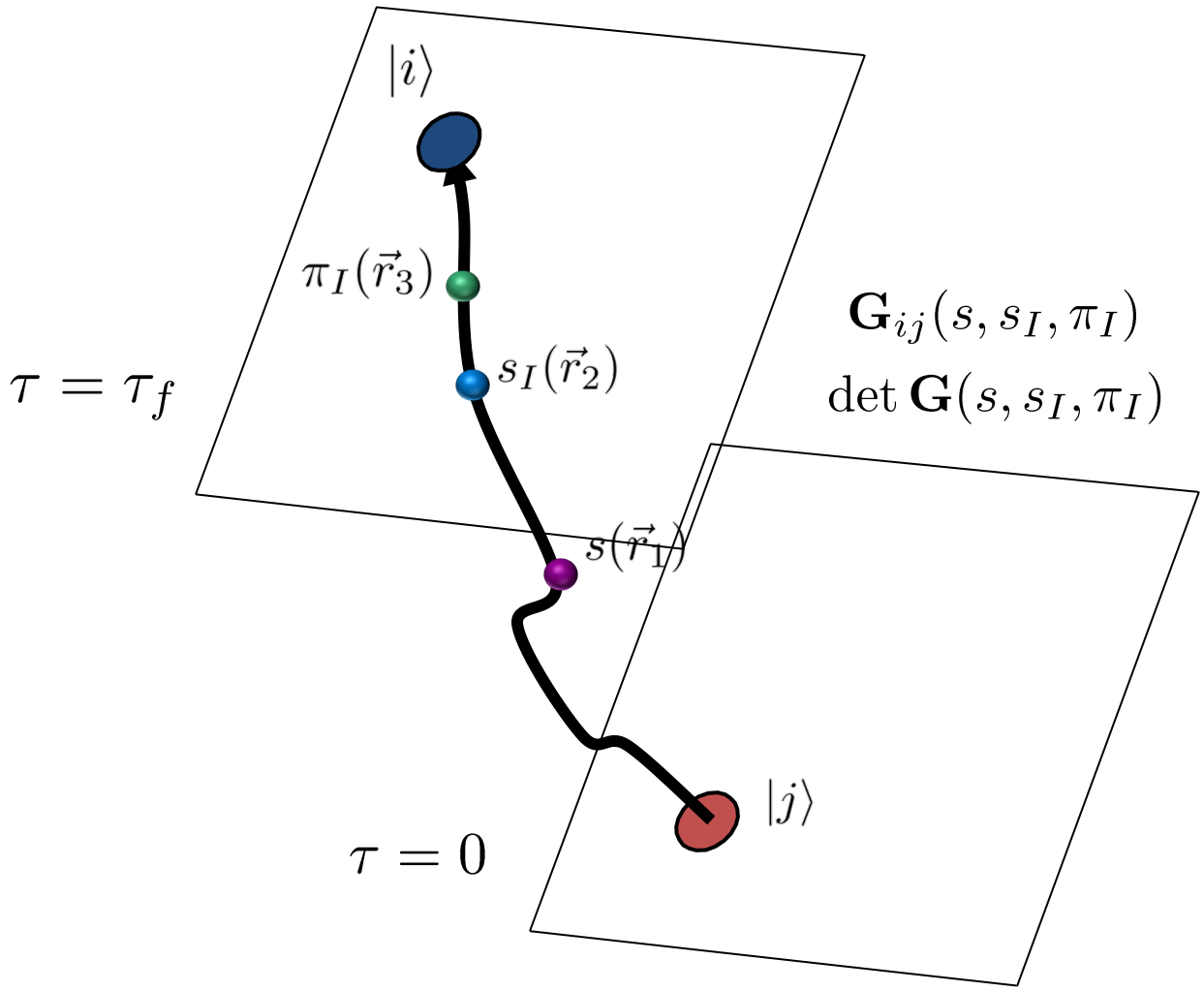
Auxiliary field method

We can write exponentials of the interaction using a Gaussian integral identity

$$\exp\left[-\frac{C}{2}(N^\dagger N)^2\right] \qquad \times \quad (N^\dagger N)^2$$

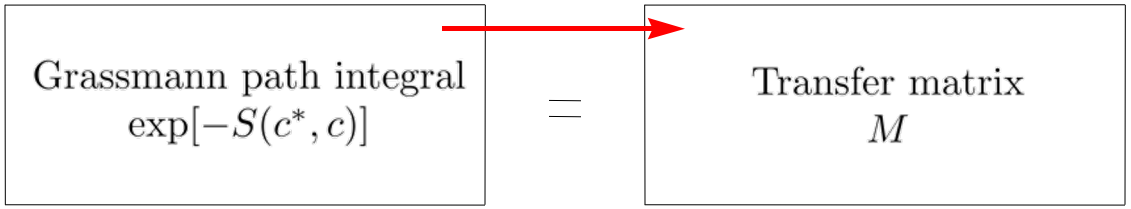
$$= \sqrt{\frac{1}{2\pi}}\int_{-\infty}^{\infty} ds \exp\left[-\frac{1}{2}s^2 + \sqrt{-C}\, s(N^\dagger N)\right] \qquad \rangle \bullet \; sN^\dagger N$$

We remove the interaction between nucleons and replace it with the interactions of each nucleon with a background field.


$|i\rangle$
$\pi_I(\vec{r}_3)$
$s_I(\vec{r}_2)$
$s(\vec{r}_1)$
$|j\rangle$
$\tau = \tau_f$
$\tau = 0$
$\mathbf{G}_{ij}(s, s_I, \pi_I)$
$\det \mathbf{G}(s, s_I, \pi_I)$

Exact equivalence of lattice formulations

We show the exact equivalence between the lattice path integrals and transfer matrix operators.



We discuss the case of fermionic particles, however the case for bosonic particles is also handled by giving the fermions fictitious labels to make them distinguishable and then symmetrizing over the fictitious labels.

For simplicity we discuss the example of two-component fermions on the lattice with contact interactions

Grassmann path integral

The path integral formulation is perhaps the most general framework for quantum fields. This is the formalism which extends rigorously to gauge fields. Convenient for the simple derivation of exact conservation laws, Noether currents, and Feynman diagram rules.

Let us consider anticommuting Grassmann fields for two-component fermions on a spacetime lattice

$$c_\uparrow(\vec{n}, n_t), c_\downarrow(\vec{n}, n_t), c_\uparrow^*(\vec{n}, n_t), c_\downarrow^*(\vec{n}, n_t)$$

The Grassmann fields are periodic in the spatial directions

$$c_i(\vec{n} + L\hat{1}, n_t) = c_i(\vec{n} + L\hat{2}, n_t) = c_i(\vec{n} + L\hat{3}, n_t) = c_i(\vec{n}, n_t)$$
$$c_i^*(\vec{n} + L\hat{1}, n_t) = c_i^*(\vec{n} + L\hat{2}, n_t) = c_i^*(\vec{n} + L\hat{3}, n_t) = c_i^*(\vec{n}, n_t)$$

and antiperiodic in the temporal direction

$$c_i(\vec{n}, n_t + L_t) = -c_i(\vec{n}, n_t)$$
$$c_i^*(\vec{n}, n_t + L_t) = -c_i^*(\vec{n}, n_t)$$

Why antiperiodic? The answer to this question is left as an exercise. We use the standard definition for the Grassmann integration

$$\int dc_i(\vec{n}, n_t) = \int dc_i^*(\vec{n}, n_t) = 0,$$
$$\int dc_i(\vec{n}, n_t) c_i(\vec{n}, n_t) = \int dc_i^*(\vec{n}, n_t) c_i^*(\vec{n}, n_t) = 1 \quad \text{(no sum on } i\text{)}$$

We note the equivalence of integration and differentiation with respect to a Grassmann variable

$$\int dc_i(\vec{n}, n_t) = \frac{\partial}{\partial c_i(\vec{n}, n_t)} \qquad \int dc_i^*(\vec{n}, n_t) = \frac{\partial}{\partial c_i^*(\vec{n}, n_t)}$$

We use the following shorthand notation for the full integration measure over all Grassmann variables

$$DcDc^* = \prod_{\vec{n},n_t,i} dc_i(\vec{n},n_t)dc_i^*(\vec{n},n_t)$$

Define the local Grassmann spin densities

$$\rho_\uparrow^{c^*,c}(\vec{n},n_t) = c_\uparrow^*(\vec{n},n_t)c_\uparrow(\vec{n},n_t),$$
$$\rho_\downarrow^{c^*,c}(\vec{n},n_t) = c_\downarrow^*(\vec{n},n_t)c_\downarrow(\vec{n},n_t),$$

and the total Grassmann density

$$\rho^{c^*,c}(\vec{n},n_t) = \rho_\uparrow^{c^*,c}(\vec{n},n_t) + \rho_\downarrow^{c^*,c}(\vec{n},n_t)$$

Define the lattice kinetic energy "hopping" coefficients

$$w_0, w_1, w_2, w_3, \cdots$$

These are defined to give a quadratic kinetic energy as function of momentum

$$w_0 - w_1 \cos q_l + w_2 \cos 2q_l - w_3 \cos 3q_l + \cdots = \frac{q_l^2}{2}\left[1 + O(q_l^{2\nu+2})\right]$$

We can take different order of lattice improvement for the kinetic energy

$$O(a^0) : \omega_0 = 1, \quad \omega_1 = 1, \quad \omega_2 = 0, \quad \omega_3 = 0$$
$$O(a^2) : \omega_0 = \tfrac{5}{4}, \quad \omega_1 = \tfrac{4}{3}, \quad \omega_2 = \tfrac{1}{12}, \quad \omega_3 = 0$$
$$O(a^4) : \omega_0 = \tfrac{49}{36}, \quad \omega_1 = \tfrac{3}{2}, \quad \omega_2 = \tfrac{3}{20}, \quad \omega_3 = \tfrac{1}{90}$$

In our simulations of nucleons, we typically use fourth-order improvement for the kinetic energy, but for illustrative simplicity we continue the discussion with the simplest case,

$$O(a^0) : \omega_0 = 1, \quad \omega_1 = 1, \quad \omega_2 = 0, \quad \omega_3 = 0$$

We use lattice units where everything is divided or multiplied by powers of the spatial lattice spacing to make it dimensionless. We also define the ratio of temporal to spatial lattice spacings

$$\alpha_t = a_t/a$$

The free nonrelativistic particle lattice action in its simplest form is

$$\to c_i^* \frac{\partial c_i}{\partial t}$$

$$S_{\rm free}(c^*,c) = \sum_{\vec{n},n_t,i} \boxed{c_i^*(\vec{n},n_t)\left[c_i(\vec{n},n_t+1) - c_i(\vec{n},n_t)\right]}$$

$$-\frac{\alpha_t}{2m}\sum_{\vec{n},n_t,i}\sum_{l=1,2,3} \boxed{c_i^*(\vec{n},n_t)\left[c_i(\vec{n}+\hat{l},n_t) - 2c_i(\vec{n},n_t) + c_i(\vec{n}-\hat{l},n_t)\right]}$$

$$\to c_i^* \frac{\partial^2 c_i}{\partial x_l^2}$$

With a contact interaction between the two components, the lattice action is

$$S(c^*, c) = S_{\text{free}}(c^*, c) + C\alpha_t \sum_{\vec{n}, n_t} \rho_{\uparrow}^{c^*, c}(\vec{n}, n_t) \rho_{\downarrow}^{c^*, c}(\vec{n}, n_t).$$

We are interested in the path integral of the exponential of the action

$$\mathcal{Z} = \int DcDc^* \exp\left[-S\left(c^*, c\right)\right]$$

Second quantization and the transfer matrix

Consider now fermion annihilation and creation operators. For the moment we consider just one operator each

$$\{a, a\} = \left\{a^\dagger, a^\dagger\right\} = 0$$
$$\left\{a, a^\dagger\right\} = 1$$

For any function of the annihilation and creation operators

$$f\left(a^\dagger, a\right)$$

we note that the quantum-mechanical trace of the normal-ordered product satisfies the following identity relating it to a Grassmann integral

$$Tr\left[:f\left(a^\dagger, a\right):\right] = \int dc dc^* e^{2c^* c} f(c^*, c)$$

Creutz, Found. Phys. 30 (2000) 487

The pedestrian proof consists of testing all four linearly independent functions of the annihilation and creation operators

$$f(a^\dagger, a) = \left\{1, a, a^\dagger, a^\dagger a\right\}$$

$$Tr\left[: f\left(a^\dagger, a\right) :\right] = \int dcdc^* e^{2c^* c} f(c^*, c)$$

$Tr\left[: f\left(a^\dagger, a\right) :\right]$

1	2
a	0
$a^\dagger$	0
$a^\dagger a$	1

$$\int dcdc^* e^{2c^* c} f(c^*, c)$$
$$= \int dcdc^* (1 + 2c^* c) f(c^*, c)$$
$$= \left(\frac{\partial}{\partial c}\right)\left(\frac{\partial}{\partial c^*}\right)(1 + 2c^* c) f(c^*, c)$$

1	2
c	0
c^*	0
$c^* c$	1

Let us rewrite the identity in a fancy form that starts to resemble the lattice Grassmann path integral

$$Tr\left[:f\left(a^{\dagger},a\right):\right]=\int dc(0)dc^{*}(0)e^{c^{*}(0)[c(0)-c(1)]}f\left[c^{*}(0),c(0)\right]$$
$$c(1)=-c(0)$$

This identity can be generalized to any sequential product of normal-ordered functions of the annihilation and creation operators.

$$Tr\left\{:f_{L_t-1}(a^{\dagger},a):\cdots:f_0(a^{\dagger},a):\right\}$$
$$=\int DcDc^{*}\exp\left\{\sum_{n_t=0}^{L_t-1}\sum_{\vec{n},i}c^{*}(n_t)\left[c(n_t)-c(n_t+1)\right]\right\}$$
$$\times f_{L_t-1}\left[c^{*}(L_t-1),c(L_t-1)\right]\cdots f_0\left[c^{*}(0),c(0)\right]$$

$$c(L_t)=-c(0)$$
$$DcDc^{*}\equiv dc(L_t-1)dc^{*}(L_t-1)\cdots dc(0)dc^{*}(0)$$

This identity can be further generalized to the case with many Grassmann variables

$$Tr\left\{:f_{L_t-1}\left[a^\dagger_{i'}(\vec{n}'),a_i(\vec{n})\right]:\cdots:f_0\left[a^\dagger_{i'}(\vec{n}'),a_i(\vec{n})\right]:\right\}$$
$$=\int DcDc^*\exp\left\{\sum_{n_t=0}^{L_t-1}\sum_{\vec{n},i}c_i^*(\vec{n},n_t)\left[c_i(\vec{n},n_t)-c_i(\vec{n},n_t+1)\right]\right\}$$
$$\times f_{L_t-1}\left[c^*_{i'}(\vec{n}',L_t-1),c_i(\vec{n},L_t-1)\right]\cdots f_0\left[c^*_{i'}(\vec{n}',0),c_i(\vec{n},0)\right]$$

with antiperiodic time boundary conditions

$$c_i(\vec{n},L_t)=-c_i(\vec{n},0)$$

D. L., Prog. Part. Nucl. Phys. 92 (2009) 117

Lähde, Meißner, "Nuclear Lattice Effective Field Theory: An Introduction", Springer (2019)

We now define the free nonrelativistic lattice Hamiltonian in its simplest form

$$H_{\text{free}} = -\frac{1}{2m}\sum_{\vec{n},i}\sum_{l=1,2,3} a_i^\dagger(\vec{n})\left[a_i(\vec{n}+\hat{l}) - 2a_i(\vec{n}) + a_i(\vec{n}-\hat{l})\right]$$

$$\rightarrow a_i^\dagger \frac{\partial^2 a_i}{\partial x_l^2}$$

We also define the following density operators

$$\rho_\uparrow(\vec{n}) = a_\uparrow^\dagger(\vec{n})a_\uparrow(\vec{n}) \qquad \rho_\downarrow(\vec{n}) = a_\downarrow^\dagger(\vec{n})a_\downarrow(\vec{n})$$

$$\rho(\vec{n}) = \rho_\uparrow(\vec{n}) + \rho_\downarrow(\vec{n})$$

So now the same Grassmann path integral we had defined before

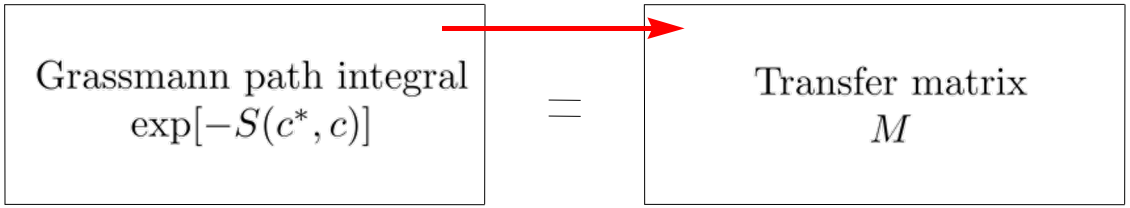
$$\mathcal{Z} = \int DcDc^* \exp\left[-S(c^*, c)\right]$$

can be rewritten in terms of the quantum-mechanical trace of the product of normal-ordered transfer matrices

$$\mathcal{Z} = Tr\left(M^{L_t}\right)$$

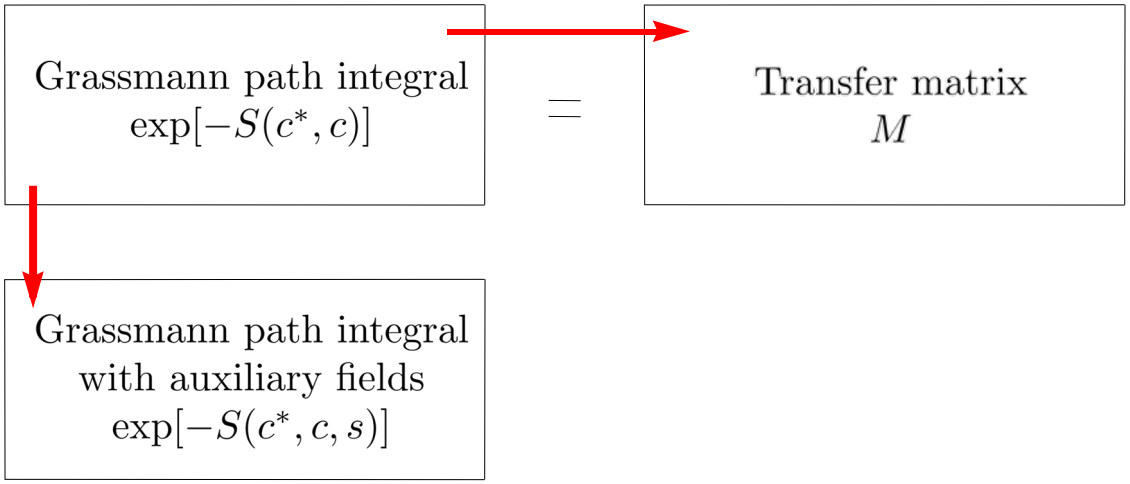
$$M =: \exp\left[-H_{\rm free}\alpha_t - C\alpha_t \textstyle\sum_{\vec{n}} \rho_\uparrow(\vec{n})\rho_\downarrow(\vec{n})\right] :$$

This demonstrates the exact equivalence of the two lattice formulations for any spatial and temporal lattice spacings.



Auxiliary fields

We now show the exact equivalence between the Grassmann path integral and the Grassmann path integral with auxiliary fields



Grassmann path integral with auxiliary fields

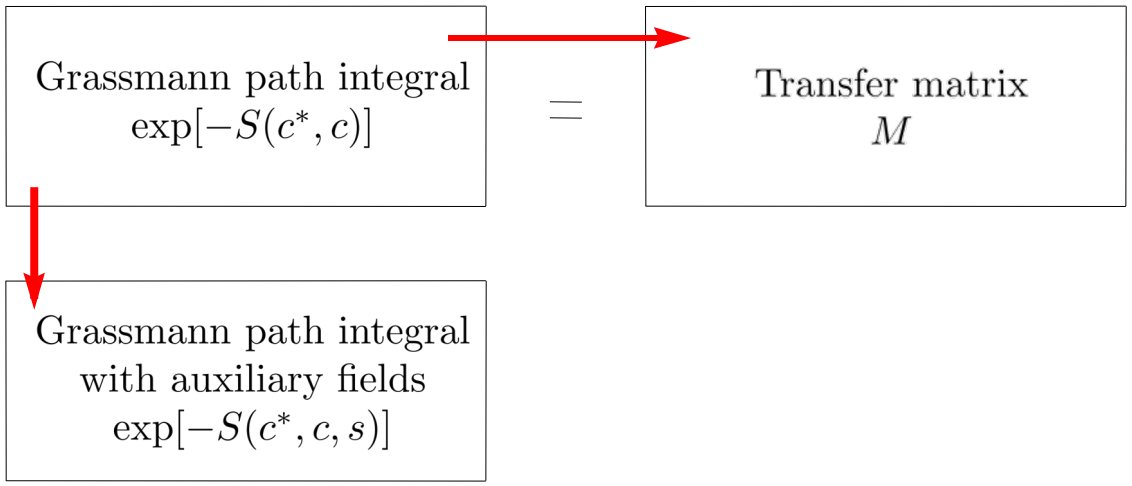
We can rewrite the same lattice Grassmann path integral using an auxiliary field

$$\mathcal{Z} = \prod_{\vec{n},n_t} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ds(\vec{n},n_t) e^{-\frac{1}{2}s^2(\vec{n},n_t)} \right] \int DcDc^* \exp\left[-S_A\left(c^*,c,s\right)\right]$$

$$S_A\left(c^*,c,s\right) = S_{\text{free}}(c^*,c) - \sum_{\vec{n},n_t} \sqrt{-C\alpha_t} s(\vec{n},n_t) \rho^{c^*,c}(\vec{n},n_t)$$

We remove the interaction between nucleons and replace it with the interactions of each nucleon with a background field.

This demonstrates the exact equivalence of the following three lattice formulations for arbitrary lattice spacings:



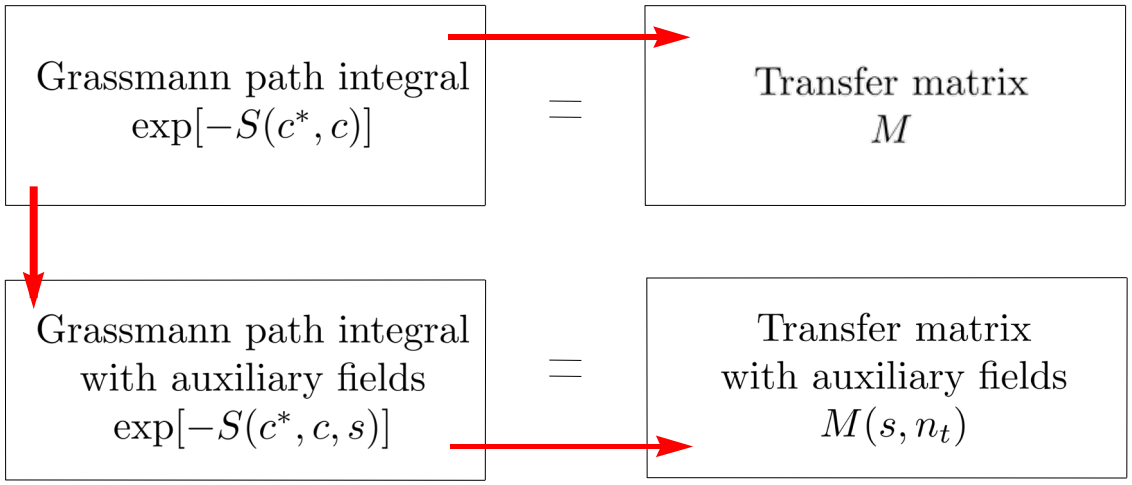
Transfer matrix operator with auxiliary fields

We use the equivalence of the Grassmann path integral and normal-ordered transfer matrix and apply it to the case of the auxiliary-field Grassmann path integral. We find

$$\mathcal{Z} = \prod_{\vec{n},n_t} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ds(\vec{n},n_t) e^{-\frac{1}{2}s^2(\vec{n},n_t)} \right] Tr\left\{ M_A(s,L_t-1) \cdot \cdots \cdot M_A(s,0) \right\}$$

$$M_A(s,n_t) =: \exp\left\{ -H_{\text{free}}\alpha_t + \sum_{\vec{n}} \sqrt{-C\alpha_t}\, s(\vec{n},n_t)\rho(\vec{n}) \right\} :$$

This shows the exact equivalence of the following four lattice formulations for arbitrary lattice spacings:



Introduction to Monte Carlo techniques

Consider approximating a one-dimensional integral by a simple Riemann sum

$$I = \int_0^L dx\, f(x) \approx \frac{L}{N} \sum_{j=1}^{N} f(x^{(j)})$$

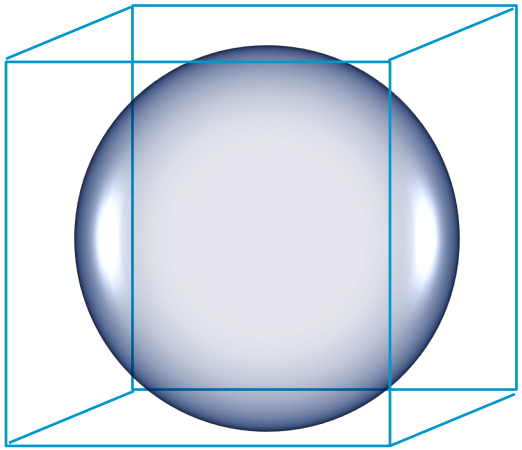
The $x^{(j)}$ are at regularly spaced intervals.

Generalization to d dimensions is more challenging. If we keep the same number of grid points, N, then only $N^{1/d}$ points per dimension and relative error can be quite large.

$$\begin{aligned} I &= \int_0^L dx_1 \cdots \int_0^L dx_d\, f(x_1, \cdots, x_d) \\ &\approx \frac{L^d}{N} \sum_{j_1=1}^{N^{1/d}} \cdots \sum_{j_d=1}^{N^{1/d}} f(x_1^{(j_1)}, \cdots, x_d^{(j_d)}) \end{aligned}$$

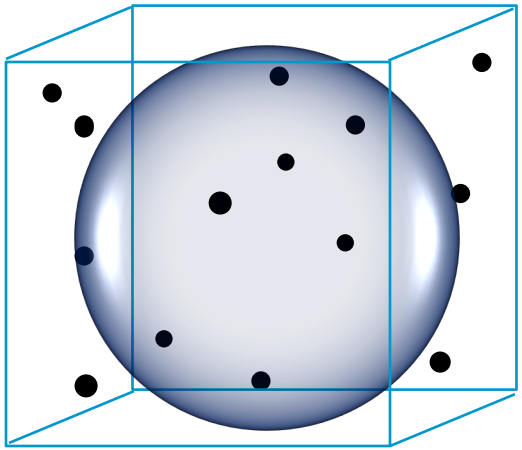
Suppose instead we choose the points at random positions in the d-dimensional space. Then relative error is purely statistical and can be as small as $N^{-1/2}$.

Example: Volume of 3D sphere using random darts



$R = 1$

$$\frac{4\pi}{3} = \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} dxdydz\,\theta(1-x^2-y^2-z^2)$$



Select N uniformly distributed random points inside the cube. Determine the fraction of points inside the sphere. This gives an estimate for the ratio of the sphere volume to cube volume.

$$\begin{aligned}\frac{4\pi}{3} &= V_{\text{cube}} \frac{V_{\text{sphere}}}{V_{\text{cube}}} \\ &\approx 8 \cdot \frac{1}{N} \sum_{i=1}^{N} \theta \left[1 - \left(x^{(i)}\right)^2 - \left(y^{(i)}\right)^2 - \left(z^{(i)}\right)^2\right]\end{aligned}$$

When calculating thermal averages in statistical mechanics or path integrals in Euclidean-time field theory, one computes sums or integrals over many degrees of freedom weighted by an exponential Boltzmann factor

$$\langle M \rangle = \frac{\sum_C M(C) e^{-\beta E(C)}}{\sum_C e^{-\beta E(C)}}$$

Due to the exponential weight, nearly all of the configurations make only a very small contribution. So a simple dartboard random sampling is very inefficient.

Importance sampling

A more efficient method to calculate the average is to select configurations with probability equal to

$$p_\beta(C) = \frac{e^{-\beta E(C)}}{\sum_{C'} e^{-\beta E(C')}}$$

This technique is called importance sampling. The thermal average is simply an average over representative configurations selected with this sampling probability.

$$\langle M \rangle = \frac{1}{N} \sum_j M(C^{(j)})$$

Markov chains

We will be discussing Markov chain algorithms, and so it is useful to review the elements and theory of Markov chains. Consider a chain of configurations labeled by order of selection. We call this integer-valued label the computation step.

Let us denote the probability of selecting configuration A at computation step n as

$$P(A, n)$$

Suppose we have selected configuration A at computation step n. The probability that we select configuration B at computation step $n + 1$ is denoted

$$W(A \to B)$$

This transition probability is chosen to be independent of n and independent of the history of configurations selected prior to selecting A at computation step n. This defines a Markov chain.

We note that

$$P(A, n+1) = P(A, n) + \sum_{B \neq A} W(B \to A) P(B, n) - \sum_{B \neq A} W(A \to B) P(A, n)$$

We now define the notion of ergodicity. Suppose we are at configuration A at computation step, n. Let S_A be the set of all positive integers m, such that the return probability to A is nonzero

$$S_A = \{m | P(A, n+m) > 0\}$$

If the set S_A is not empty, then we say that A is positive recurrent. If the greatest common divisor of the set of integers in S_A is 1, then we say that A is aperiodic. If all of the configurations connected by the Markov chain are recurrent and aperiodic, then the Markov chain is said to be ergodic. If the Markov chain is ergodic and all configurations are connected by the graph of nonzero transitions in the Markov chain, then there is a unique equilibrium distribution that is reached in the limit of large number of computation steps that is independent of the initial conditions.

$$\lim_{\tau \to \infty} P(C, \tau) \to p(C)$$

Serfozo, "Basics of Applied Stochastic Processes", (Berlin: Springer-Verlag) 2009

Detailed balance

We want the equilibrium probability distribution to be

$$p_{\text{target}}(C)$$

One way to do this is to require

$$W(A \to B)p_{\text{target}}(A) = W(B \to A)p_{\text{target}}(B)$$

for every pair of configurations A and B. This condition is called detailed balance.

If the Markov chain is ergodic and all configurations are connected, then after many computation steps we reach the unique equilibrium distribution, which satisfies the stationary condition

$$\sum_{B \neq A} W(A \to B)p(A) = \sum_{B \neq A} W(B \to A)p(B)$$

Comparing with the detailed balance condition, we conclude that

$$p(A) = p_{\text{target}}(A)$$

for all configurations A.

Metropolis algorithm

One popular method for generating the desired detailed balance condition is the Metropolis algorithm

Metropolis, Teller, Rosenbluth, J. Chem. Phys. 21 (1953) 1087

$$W(A \to B) = \begin{cases} \frac{p_{\text{target}}(B)}{p_{\text{target}}(A)} & p_{\text{target}}(B) \leq p_{\text{target}}(A) \\ 1 & p_{\text{target}}(B) > p_{\text{target}}(A) \end{cases}$$

Usually the transition probability can be divided in terms of a proposed move probability and an acceptance probability,

$$W(A \to B) = W_{\text{propose}}(A \to B) W_{\text{accept}}(A \to B)$$

And quite often the proposed move probability is symmetric

$$W_{\text{propose}}(A \to B) = W_{\text{propose}}(B \to A)$$

However this does not need to be the case. One can design useful algorithms where there is some guiding involved in the proposed moves. It is also not necessary that you use only one type of update. If you maintain detailed balance for each type of update process, then you also recover the target probability distribution.

Once your Markov chain is set up properly, you can now compute observables such as

$$\langle O \rangle = \frac{\sum_A O(A) p_{\text{target}}(A)}{\sum_A p_{\text{target}}(A)}$$

by computing the average

$$\langle O \rangle = \frac{\sum_{n=1,N} O(A_n)}{N}$$

for large N from your Markov chain.

Review of topics

Introduction and history

Grassmann path integrals

Transfer matrix operators

Auxiliary fields

Monte Carlo methods

Markov chain Monte Carlo