
An introduction to lattice QCD with and without AI

P2EAK 2026 講義ノート

<https://indico.yukawa.kyoto-u.ac.jp/event/98/>

Akio Tomiya

Tokyo Woman's Christian University

akio@yukawa.kyoto-u.ac.jp

2026 年 7 月 11 日

概要

本ノートでは、場の量子論を有限自由度の確率分布として定義する方法として格子場理論を導入し、その分布から物理を変えずにサンプルする方法としてハイブリッド・モンテカルロ法（HMC 法）と正規化フローを説明する。第 1 講ではユークリッド経路積分、実スカラー場、SU(3) 格子ゲージ理論、格子フェルミオンを扱う。第 2 講ではマルコフ連鎖、自己相関、HMC 法、正規化フロー、Metropolis–Hastings 補正を扱う。第 3 講はおまけのハンズオンとして、LatticeQCD.jl を使い、クエンチ配位の生成からパイオン相関関数、cosh 型有効質量、block Jackknife 誤差までを実行する。式の操作だけでなく、各構成要素が何を保証し、どこで計算が難しくなるかを図と具体例で示す。

目次

学習目標と記号	1
第 1 講: 格子場理論が定める確率分布	2
1 ユークリッド場の理論へ	2
1.1 時間に依存しないハミルトニアン	2
1.2 虚時間相関関数から同じスペクトルが得られる理由	2
1.3 演算子表示から Euclid 経路積分へ: 1 自由度系の Trotter 分解	3
1.4 格子場の理論は場の理論の 1 つの定義	4
2 実スカラー場による最小例	6
2.1 連続理論と格子作用	6
2.2 古典連続極限と量子連続極限	7
2.3 量子連続極限とくりこみ条件	8
2.4 Wilson 型くりこみ群とユニバーサリティ	9
3 SU(3) 格子ゲージ理論	9
3.1 $U(1)$ ゲージ対称性と Maxwell–Dirac 理論	9
3.2 Yang–Mills 理論と内山の一般化	10
3.3 平行移動子によるゲージ共変差分	11
3.4 ゲージ対称性を保った格子ゲージ作用	12
3.5 最小閉路としてのプラケット	12
3.6 Wilson 作用	12
3.7 プラケットと連続ゲージ作用	13
4 格子フェルミオン	14
4.1 ナイーブ差分とダブリング	14
4.2 Nielsen–Ninomiya 定理	15
4.3 Wilson フェルミオン	16
5 第 1 講のまとめ	17
第 2 講: 格子場理論におけるサンプリングと機械学習	18

6	モンテカルロ積分とマルコフ連鎖	18
6.1	台形則・区分求積と次元の呪い	18
6.2	独立標本による期待値の推定	19
6.3	2点相関関数と質量の測定	20
6.4	重要度サンプリングとマルコフ連鎖 Monte Carlo 法	21
6.5	Glasgow の天気模型と確率過程の収束	21
6.6	Metropolis–Hastings 法	22
6.7	自己相関と有効サンプル数	23
7	ハイブリッド・モンテカルロ法	24
7.1	仮想的なハミルトン力学	24
7.2	HMC 法の 1 更新	26
7.3	自由スカラー場の HMC 力	26
8	新しいサンプリングの必要性	27
8.1	HMC 法の計算コストと臨界減速	27
8.2	全結合ニューラルネット	28
8.3	学習モデルと厳密性の役割分担	29
9	正規化フローと Metropolis 補正	30
9.1	変数変換に基づく生成モデル	30
9.2	逆 KL 学習と Metropolis–Hastings 補正	31
9.3	研究例：ゲージ同変フローとトポロジーの凍結	33
9.4	研究例：動的クォークを含む 4 次元 QCD	33
10	拡散模型	34
10.1	研究例：Lingxiao Wang らの拡散サンプラー	35
10.2	研究例：自己学習拡散	36
11	第 2 講のまとめ	37
第 3 講: おまけ: LatticeQCD.jl でパイオン質量を測る		38
12	この実習で何を学ぶか	38
13	パイオン質量を測るとはということか	39
14	Euclid 経路積分から Monte Carlo 法へ	40
15	他の格子 QCD コードと JuliaQCD の位置づけ	41
16	JuliaQCD を構成するパッケージ	42
17	再現可能な二つの実行経路	43
18	Windows、macOS、Linux への Julia 導入	44
19	Project と Manifest を固定する	45

20	Google Colab: Julia を毎回入れ直さない	46
21	使用する物理・数値パラメータ	47
22	Heatbath 更新、overrelaxation、熱化	48
23	Wilson フェルミオンと hopping parameter	49
24	パイオン演算子と伝播関数の縮約	50
25	周期格子の相関関数と cosh	51
26	実行に用いる TOML 設定ファイル	52
27	シミュレーションを実行し、ログを読む	53
28	未加工の相関関数を厳密に読む	54
29	平均、時間反転、相関関数の形	55
30	cosh 型有効質量と plateau	56
31	Jackknife 法と JkTools: 非線形解析を丸ごと再計算	57
32	最終結果、限界、次の演習	58
	補足資料	59
A1	くりこみ群変換で何が不変になるか	59
A2	虚時間相関関数と質量	59
A3	Euclid 相関関数と再構成定理	59
A4	差分演算子の古典連続極限	60
A5	積分測度、自由度とエリツツァーの定理	61
A6	代表的なゲージ不変観測量	62
A7	Wilson 作用による数値計算	63
A8	グラスマン積分とフェルミオン行列式	63
A9	擬フェルミオン表示	64
A10	重要度サンプリングとマルコフ連鎖 Monte Carlo 法	65
A11	局所ランダムウォークの限界	65
A12	HMC 法が目標分布を保つ理由	65

A13	ゲージ場に対する HMC 法	66
A14	コードとの対応	66
A15	ニューラルネットを用いた確率分布の表現	67
A16	ヤコビアン計算の課題と設計上の工夫	67
A17	正規化フローの発展と格子場理論への応用	68
A18	1 変数変換の幾何学的理解	68
A19	ゲージ共変フロー	68
A20	HMC 法との組み合わせ	69
A21	サンプラーの評価基準	70
A22	絶対零度と松原形式	71
A23	スペクトル表示と実時間再構成	71
A24	自由場と多変量ガウス分布	71
A25	自由場の運動量表示と格子伝播関数	72
A26	Gell–Mann–Low 方程式	73
A27	格子上の摂動論	74
A28	ダブラーが連続極限で消えない理由	74
A29	point splitting 法とゲージ不変な複合演算子	75
A30	マルコフ連鎖の定常性とエルゴード性	75
A31	Metropolis–Hastings 受容率と詳細釣り合い	76
A32	学習モデルと厳密性の役割分担	76
A33	リープフロッグ積分	77
A34	シンプレクティック積分とエネルギー誤差	77
A35	2 変数スカラー場に対する HMC 法の 1 更新	78
	A35.1 2 変数モデルから分かること	79
	A35.2 ガウス分布における HMC 軌道	79
A36	ブラケットの連続極限	79
	A36.1 可換な $U(1)$ 理論との比較	80
	A36.2 トレースと連続極限	80

A37	Polyakov loop	80
A38	アフィンカップリング層とヤコビアン の計算	81
A39	1 次元フローと MH 補正	82
	A39.1 線形なスケール変換	82
	A39.2 非線形変換と確率密度	82
	A39.3 不完全な提案分布に対する MH 補正	83
	A39.4 確率分布の台に関する条件	83
	A39.5 受容率と重みのばらつき	83
A40	モード崩壊と分布の被覆	83
A41	有限時間幅と有効質量	84
A42	重要公式一覧	84

学習目標と記号

本書の主題は、場の量子論が定める高次元確率分布から、物理を変えずにサンプルする方法である。前半ではこの確率分布の定義を、後半ではそこからのサンプル法を扱う。

到達目標

この講義の到達目標は次の3点である。

1. 格子化によって場の量子論がどのような有限次元分布になるか。
2. HMC 法が高次元分布を探索しつつ、なぜ厳密な期待値を与えるか。
3. 正規化フローを使うとき、学習誤差と物理的な偏りをどう分離するか。
4. LatticeQCD.jl でパイオン相関関数を生成し、質量と統計誤差を求める手順。

記号と前提

自然単位系 $\hbar = c = 1$ を用いる。特に断らない限り時空は4次元 Euclid 空間、格子間隔は a 、格子点は n 、正方向単位ベクトルは $\hat{\mu}$ と書く。

記号	意味
S_M, S_E	Minkowski 作用と Euclid 作用
ϕ_n, φ_n	格子点 n 上の次元付き・無次元スカラー場
$U_\mu(n)$	n から $n + \hat{\mu}$ へ向く SU(3) リンク変数
$D[U]$	ゲージ場 U 上の格子 Dirac 演算子
$\pi(x)$	物理が定める目標分布。円周率との混同を避けるため文脈を明示する
$q_\theta(x)$	学習した提案分布またはモデル分布
τ	Euclid 時間または HMC の仮想時間。節ごとに意味を明示する

第1講

格子場理論が定める確率分布

1 ユークリッド場の理論へ

1.1 時間に依存しないハミルトニアン

この節では、ハミルトニアン H が時刻 t に陽に依存しない系を考える。Schrödinger 方程式

$$i \frac{d}{dt} |\Psi(t)\rangle = H |\Psi(t)\rangle \quad (1)$$

の解は

$$|\Psi(t)\rangle = e^{-iHt} |\Psi(0)\rangle \quad (2)$$

である。 $H|n\rangle = E_n|n\rangle$ とすると、各エネルギー固有状態は

$$|n; t\rangle = e^{-iE_n t} |n\rangle \quad (3)$$

と発展する。実時間では、エネルギー E_n は位相の回転速度として現れる。

1.2 虚時間相関関数から同じスペクトルが得られる理由

虚時間 τ を

$$t = -i\tau \quad (4)$$

で導入する。この実時間から虚時間への変換を Wick 回転と呼ぶ。すると、式 (2) は

$$|\Psi(\tau)\rangle = e^{-H\tau} |\Psi(0)\rangle, \quad |n; \tau\rangle = e^{-E_n \tau} |n\rangle \quad (5)$$

となる。ハミルトニアンは置き換わっておらず、時間に依存しない同じ H を用いているため、固有状態 $|n\rangle$ も固有値 E_n も実時間と虚時間で共通である。¹ 変わるのは、 $e^{-iE_n t}$ という振動が $e^{-E_n \tau}$ という減衰になることだけである。

実時間と虚時間でスペクトルが共通する理由

H が t に依存しないので、時間発展演算子は H の通常の指数関数として書ける。Wick 回転は時間変数を複素方向へ動かすが、 H の固有値問題 $H|n\rangle = E_n|n\rangle$ は変えない。そのため虚時間相関関数の指数減衰率から、実時間理論と同じエネルギースペクトルを取り出せる。

¹もし $H = H(t)$ が時刻に陽に依存するなら、 $U(t, 0) = \mathcal{T} \exp \left[-i \int_0^t dt' H(t') \right]$ のような時間順序積が必要になる。この場合は、単一の時間非依存 H の固有値を実時間と虚時間で共有するという本文の単純な説明は、そのままでは使えない。

1.3 演算子表示から Euclid 経路積分へ：1 自由度系の Trotter 分解

統計力学の分配関数も同じ演算子 H から

$$Z(\beta) = \text{Tr} e^{-\beta H} = \sum_n e^{-\beta E_n} \quad (6)$$

と書かれる。 $e^{-\beta H}$ を短い虚時間幅 $\epsilon = \beta/N_\tau$ へ分解し、各時刻の間に場の固有状態の完全系を挿入すると経路積分が得られる。この計算を 1 自由度で具体的に追う。

ハミルトニアン

$$H = \frac{p^2}{2M} + V(q) \quad (7)$$

を考える。 $\epsilon = \beta/N_\tau$ として

$$e^{-\beta H} = (e^{-\epsilon H})^{N_\tau} \quad (8)$$

と分割し、各区間に

$$1 = \int dq_k |q_k\rangle\langle q_k|, \quad 1 = \int \frac{dp_k}{2\pi} |p_k\rangle\langle p_k| \quad (9)$$

を挿入する。 $[p^2, V(q)] \neq 0$ だが、短時間では Trotter 公式により

$$e^{-\epsilon H} = e^{-\epsilon p^2/(2M)} e^{-\epsilon V(q)} + O(\epsilon^2) \quad (10)$$

と書ける。 $\langle q|p\rangle = e^{ipq}$ という規約を用いると

$$\langle q_{k+1}|e^{-\epsilon H}|q_k\rangle \simeq \int \frac{dp_k}{2\pi} \exp \left[ip_k(q_{k+1} - q_k) - \epsilon \frac{p_k^2}{2M} - \epsilon V(q_k) \right]. \quad (11)$$

運動量について平方完成すると

$$-\epsilon \frac{p_k^2}{2M} + ip_k \Delta q_k = -\frac{\epsilon}{2M} \left(p_k - iM \frac{\Delta q_k}{\epsilon} \right)^2 - \frac{M(\Delta q_k)^2}{2\epsilon}, \quad (12)$$

$\Delta q_k = q_{k+1} - q_k$ である。ガウス積分を実行して

$$\langle q_{k+1}|e^{-\epsilon H}|q_k\rangle \simeq \left(\frac{M}{2\pi\epsilon} \right)^{1/2} \exp \left[-\epsilon \left\{ \frac{M}{2} \left(\frac{q_{k+1} - q_k}{\epsilon} \right)^2 + V(q_k) \right\} \right] \quad (13)$$

を得る。トレースでは $q_{N_\tau} = q_0$ と同一視されるため

$$Z \propto \int \prod_{k=0}^{N_\tau-1} dq_k e^{-S_E^{\text{disc}}[q]}, \quad (14)$$

$$S_E^{\text{disc}}[q] = \epsilon \sum_{k=0}^{N_\tau-1} \left[\frac{M}{2} \left(\frac{q_{k+1} - q_k}{\epsilon} \right)^2 + V(q_k) \right]. \quad (15)$$

$\epsilon \rightarrow 0$ で

$$S_E[q] = \int_0^\beta d\tau \left[\frac{M}{2} \dot{q}^2 + V(q) \right] \quad (16)$$

となる。場の理論では空間点ごとに q を置いたものと考えればよく、運動項は時間差分と空間差分を合わせた $(\partial_\mu \phi)^2$ になる。したがって

$$Z = \int \mathcal{D}\phi e^{-S_E[\phi]} \quad (17)$$

という Euclid 経路積分が得られる。一方、Minkowski 時空では形式的に

$$\langle O \rangle = \frac{1}{Z_M} \int \mathcal{D}\phi O[\phi] e^{iS_M[\phi]} \quad (18)$$

である。Wick 回転により

$$iS_M \longrightarrow -S_E \quad (19)$$

となり、振動する重みが減衰する重みへ移る。

S_E が実で下に有界なら

$$p[\phi] = \frac{1}{Z} e^{-S_E[\phi]} \quad (20)$$

は確率分布と同じ形をしている。ここが格子場理論、統計力学、機械学習をつなぐ点である。

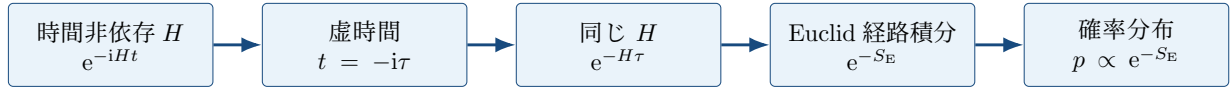


図 1: 時間非依存ハミルトニアンから Euclid 経路積分へ至る流れ。同じ H を使うためスペクトルは共通であり、虚時間ではエネルギー固有値が指数減衰率になる。

確率分布として扱えない場合

S_E が複素数なら e^{-S_E} は正の確率密度にならない。有限密度 QCD などの符号問題は、サンプラーの改良だけでは解決しない。

有限温度における虚時間の周期性、松原周波数、および絶対零度で真空情報が得られる理由は付録 A22 にまとめる。

1.4 格子場の理論は場の理論の 1 つの定義

ここまでで、Euclid 経路積分の重み $e^{-S_E[\phi]}$ が確率分布の形を持つことを見た。実際の計算では、相関関数や質量などの物理量は典型的に

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}\phi O[\phi] e^{-S_E[\phi]}, \quad Z = \int \mathcal{D}\phi e^{-S_E[\phi]} \quad (21)$$

の形で与えられる。1 粒子の量子力学なら経路 $q(\tau)$ 全体について積分する。場の量子論では、各時空点 x ごとに場 $\phi(x)$ があるので、積分変数は関数 $\phi(x)$ そのものである。記号的には

$$\mathcal{D}\phi \stackrel{\text{formal}}{\sim} \prod_{x \in \mathbb{R}^4} d\phi(x) \quad (22)$$

と書きたくなる。ここでは無限 Euclid 時空を考えて $x \in \mathbb{R}^4$ とした。有限体積なら $\mathbb{R}^4$ をその領域に置き換えるが、連続時空に無限個の点があるという問題は残る。したがって場の量子論の計算は、素朴には無限次元の積分である。

式 (22) は積分の定義ではない

問題は、単に積分変数が多くて計算が重いことではない。 $\prod_{x \in \mathbb{R}^4} d\phi(x)$ に対応する、有限次元の Lebesgue 測度をそのまま拡張した平行移動不変な測度は、無限次元の場の空間には存在しない。したがって正則化を入れる前には、積分順序の交換、変数変換、部分積分、規格化などを通常の多重積分と同じように行ってよいか明らかでなく、式 (22) だけでは具体的な操作方法が定まっていない。

自由場では、作用を $S_E = \frac{1}{2}(\phi, K\phi)$ と書き、Fourier モードごとに有限次元のガウス積分と同じ

計算を形式的に行って

$$Z_0^{\text{formal}} (\det K)^{-1/2} = \prod_p \lambda_p^{-1/2}, \quad \langle \phi(x) \phi(y) \rangle_0 = K^{-1}(x, y)$$

と書ける。しかし $\prod_p \lambda_p^{-1/2}$ 自体は一般に発散し、行列式や相関関数へ意味を与えるにはカットオフ、有限体積、格子などの正則化と、その除去の手続きが必要である。相互作用がある場合にはガウス積分への帰着もできない。つまり、正則化前の経路積分は有用な形式記号ではあるが、それだけで定義された積分ではない。

この見方に立つと、格子場理論の手続きは3段階に整理できる。まず連続時空を有限個の格子点で置き換えれば、 $\prod_{x \in \mathbb{R}^4} d\phi(x)$ という形式記号は

$$\prod_{n=1}^{V/a^4} d\phi_n \quad (23)$$

という有限次元の多重積分になる。次に、その有限次元積分を Monte Carlo 法で評価する。最後に、体積を大きくし、格子間隔を小さくする極限を調べる。つまり格子化は、無限次元積分を有限次元積分の列として定義し直す手続きである。

計算問題としての場の量子論

場の量子論で欲しいのは、場の配位 $\phi(x)$ 全体にわたる期待値である。格子化すると、この問題は巨大だが有限次元の確率分布 $p(\phi_1, \dots, \phi_N) \propto e^{-S_{\text{lat}}}$ からのサンプリング問題になる。

ここでいう離散化は、数値計算のための便宜というより、連続量に対する数学的な操作を有限個の量からなる列の極限として明文化するための定義そのものである。たとえば1変数関数 $f(x)$ の微分を定義するとき、

$$f'(x) = \lim_{a \rightarrow 0} \frac{f(x+a) - f(x)}{a} \quad (24)$$

と書ける。右辺は、有限の a では差分であり、極限を取ることで微分になる。同様に、1変数積分も

$$\int_{x_i}^{x_f} dx f(x) = \lim_{a \rightarrow 0} a \sum_{n=0}^{N-1} f(x_i + na), \quad Na = x_f - x_i \quad (25)$$

のように、有限個の和から出発して極限で定義できる。場の量子論でも同じ発想を使う。まず連続時空を離散化し、有限個の変数に対する積分として理論を定義する。その後、物理量を固定しながら連続極限を取る。数値計算は、この有限格子上の理論を評価する方法の1つであり、離散化の目的そのものではない。

間隔 a の4次元格子

$$x_\mu = an_\mu, \quad n_\mu = 0, 1, \dots, N_\mu - 1 \quad (26)$$

を導入する。有限体積では自由度の数が有限になり、汎関数積分は通常の多重積分になる。実スカラー場なら

$$\mathcal{D}\phi \longrightarrow \prod_n d\phi_n. \quad (27)$$

格子化には、紫外発散を制御する正則化と、理論を有限自由度の積分として与える非摂動的定義という2つの役割がある。有限な格子間隔と有限体積を保つ限り、適切に定義された作用に対する経路積分は有限次元積分となり、場の理論で典型的に現れる紫外発散は生じないため、理論の定義に曖昧さ

がない。したがって、数値計算を目的としない場合でも、格子正則化された場の理論を学ぶことには、場の量子論が非摂動的にどのように定義されるかを具体的に理解する意義がある。

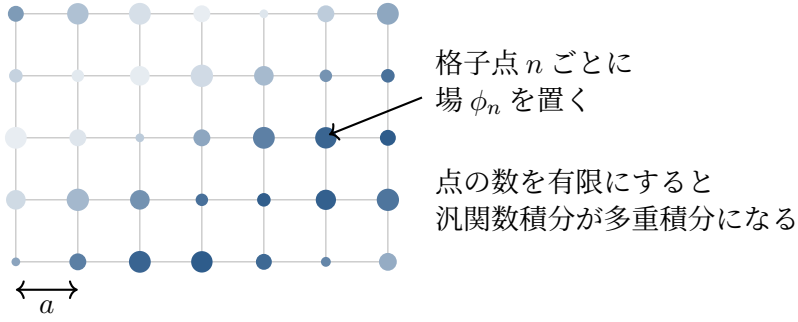


図 2: 4次元格子上の実スカラー場を2次元断面で模式的に表したもの。円の大きさと濃さが場の値を表す。

場の量子論から確率分布までの要点

虚時間形式は別のスペクトルを持つ理論ではない。時間に依存しない同じ H を用いるため、実時間での位相回転 $e^{-iE_n t}$ が虚時間での指数減衰 $e^{-E_n \tau}$ に変わり、同じ E_n を相関関数から読める。格子化は、その Euclid 経路積分を有限次元の確率分布として定義する段階である。

2 実スカラー場による最小例

2.1 連続理論と格子作用

離散時空での場の理論を考えるため、まず連続な4次元ユークリッド時空の ϕ^4 理論を考える。

$$S_E[\phi] = \int d^4x \left[\frac{1}{2} \sum_{\mu=1}^4 (\partial_\mu \phi)^2 + \frac{m_0^2}{2} \phi^2 + \frac{\lambda_0}{4!} \phi^4 \right]. \quad (28)$$

離散時空を考えるには、連続時空の微分を差分に、積分を格子点上の和に置き換えればよい。ここでは前進差分を用いて

$$\partial_\mu \phi(x) \longrightarrow \frac{\phi_{n+\hat{\mu}} - \phi_n}{a}, \quad \int d^4x \longrightarrow a^4 \sum_n \quad (29)$$

を使えば

$$S_{\text{lat}}[\phi] = \sum_n \left[\frac{a^2}{2} \sum_{\mu} (\phi_{n+\hat{\mu}} - \phi_n)^2 + \frac{a^4 m_0^2}{2} \phi_n^2 + \frac{a^4 \lambda_0}{4!} \phi_n^4 \right]. \quad (30)$$

場を $\varphi_n = a\phi_n$ と無次元化すると

$$S_{\text{lat}}[\varphi] = \sum_n \left[\frac{1}{2} \sum_{\mu} (\varphi_{n+\hat{\mu}} - \varphi_n)^2 + \frac{(am_0)^2}{2} \varphi_n^2 + \frac{\lambda_0}{4!} \varphi_n^4 \right]. \quad (31)$$

式 (31) が、無次元変数で書いた格子作用である。数値コードが直接見るのは a や m_0 ではなく、 am_0 のような無次元量である。有限格子 Λ_{lat} 上では、これに対応する経路積分測度は有限個の通常の積分

$$\int \mathcal{D}\varphi \equiv \prod_{n \in \Lambda_{\text{lat}}} \int_{-\infty}^{\infty} d\varphi_n, \quad Z_{\text{lat}} = \int \mathcal{D}\varphi e^{-S_{\text{lat}}[\varphi]}, \quad \langle O \rangle = \frac{1}{Z_{\text{lat}}} \int \mathcal{D}\varphi O[\varphi] e^{-S_{\text{lat}}[\varphi]} \quad (32)$$

である。 N_{site} を格子点数とすると、 $\varphi_n = a\phi_n$ によって元の測度は $\prod_n d\phi_n = a^{-N_{\text{site}}} \prod_n d\varphi_n$ と変わるが、この因子は配位に依存しないため Z_{lat} の規格化へ吸収され、期待値には影響しない。

Ising 模型との対応

式 (31) は連続値の場合 $\varphi_n \in \mathbb{R}$ に対する Boltzmann 重み $p(\varphi) \propto e^{-S_{\text{lat}}[\varphi]}$ を定める。一方、Ising 模型では各格子点に離散値のスピン $s_i = \pm 1$ を置く。まず隣接点の関係が見やすい 1 次元鎖を考え、周期境界条件 $s_{N+1} = s_1$ のもとでハミルトニアンを

$$\mathcal{H}_{\text{Ising}}(\{s\}) = -J \sum_{i=1}^N s_i s_{i+1} - h \sum_{i=1}^N s_i, \quad s_i = \pm 1 \quad (33)$$

と書く。統計力学の重みは

$$p(\{s\}) \propto e^{-\beta \mathcal{H}_{\text{Ising}}} = \exp \left[\beta J \sum_{i=1}^N s_i s_{i+1} + \beta h \sum_{i=1}^N s_i \right]. \quad (34)$$

ここで $s_i s_{i+1}$ は隣り合う 2 点のスピンを直接結ぶホッピング項である。2 次元以上でも同様に、各格子点から各方向の隣接点へ向かう積 $s_n s_{n+\hat{\mu}}$ を足し合わせればよい。したがって場の理論の作用 S は、統計力学でいう $\beta \mathcal{H}$ に対応する。

隣接項を比べるため、周期境界条件で式 (31) の差分項を展開すると

$$\frac{1}{2} \sum_{n,\mu} (\varphi_{n+\hat{\mu}} - \varphi_n)^2 = d \sum_n \varphi_n^2 - \sum_{n,\mu} \varphi_n \varphi_{n+\hat{\mu}}. \quad (35)$$

最後の項は、隣り合う場の値を同じ向きにそろえるので、強磁性的 Ising 模型の $J > 0$ の項に似ている。違いは、Ising 模型の変数が $s_n = \pm 1$ に固定されているのに対し、 ϕ^4 格子場では φ_n の大きさも連続的に揺らぐ点である。 $m_0^2 < 0$ で二重井戸ポテンシャルが深くなると、 φ_n は 2 つの極小付近に集まり、おおまかに $\varphi_n \simeq v s_n$ と見なせる。この意味で、Ising 模型は Z_2 対称性をもつスカラー場理論の極限的な模型として理解できる。外場 h は、スカラー場側では $\sum_n \varphi_n$ に比例する線形項を入れることに対応する。

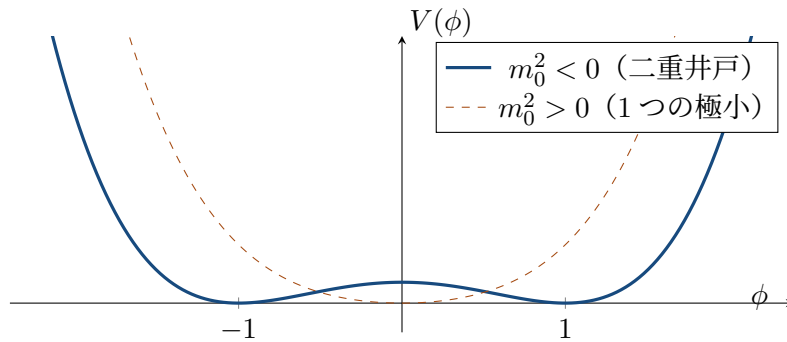


図 3: ϕ^4 ポテンシャルの典型形。隣接結合を加えると、各格子点の場は独立ではなく、空間的に相関した高次元分布を作る。

2.2 古典連続極限と量子連続極限

格子作用から連続理論を得るときには、古典的な差分の極限と、量子論としての連続極限を分けて考える必要がある。2 つを区別する。

古典連続極限 差分が微分に戻り、格子作用が連続作用に戻ることを。

量子連続極限 相関長 ξ を格子単位で無限大にし、物理相関長 $\xi_{\text{phys}} = a\xi$ を有限に保つこと。

量子連続極限は臨界点への接近として現れる。

$$\frac{a}{\xi_{\text{phys}}} = \frac{1}{\xi} \longrightarrow 0. \quad (36)$$

裸パラメータを調整しながらこの極限を取る方法は、節 2.3 で説明する。

自由作用と多変量ガウス分布の関係は付録 A24 にまとめる。

自由場の運動量表示、格子伝播関数、および格子上の摂動論への接続は、付録 A25 にまとめる。

2.3 量子連続極限とくりこみ条件

ここでは、第 2 節で扱った 4 次元 $\lambda\phi^4$ 理論を念頭に置く。作用が裸の質量 m 、裸の 4 点結合 λ 、紫外カットオフ Λ に依存するとし、期待値を

$$\langle O \rangle_{\Lambda, m, \lambda} = \frac{1}{Z_{\Lambda, m, \lambda}} \int_{|p| \lesssim \Lambda} \mathcal{D}\phi O[\phi] e^{-S_{\Lambda}[\phi; m, \lambda]} \quad (37)$$

と書く。ここで $\Lambda \simeq 1/a$ である。 m と λ でそれぞれ微分すると

$$\begin{aligned} \frac{\partial}{\partial m} \langle O \rangle_{\Lambda, m, \lambda} &= \left\langle \frac{\partial O}{\partial m} \right\rangle_{\Lambda, m, \lambda} - \left\langle O \frac{\partial S_{\Lambda}}{\partial m} \right\rangle_{c, \Lambda, m, \lambda}, \\ \frac{\partial}{\partial \lambda} \langle O \rangle_{\Lambda, m, \lambda} &= \left\langle \frac{\partial O}{\partial \lambda} \right\rangle_{\Lambda, m, \lambda} - \left\langle O \frac{\partial S_{\Lambda}}{\partial \lambda} \right\rangle_{c, \Lambda, m, \lambda}. \end{aligned} \quad (38)$$

最後の添字 c は連結部分を表す。質量を動かせば $\partial S/\partial m$ が、4 点結合を動かせば $\partial S/\partial \lambda$ が挿入された相関関数が現れ、一般に $\langle O \rangle$ は変化する。

一方、 m と λ を固定したまま Λ を動かしても期待値は変わる。積分する Fourier 成分の範囲が変わり、それらを積分消去して得られる作用の係数も変わるからである。したがって、有限カットオフで得た数値そのものを物理量とみなすことはできない。そこで、基準に選んだ相関関数 $\langle O_i \rangle$ を不変に保つように $m = m(\Lambda)$ と $\lambda = \lambda(\Lambda)$ を同時に動かすことを考える。ここで $d/d \ln \Lambda$ は、カットオフへの明示的な依存性だけでなく、 $m(\Lambda)$ と $\lambda(\Lambda)$ の変化も含む全微分である。不変性の条件は

$$\begin{aligned} 0 &= \frac{d}{d \ln \Lambda} \langle O_i \rangle_{\Lambda, m(\Lambda), \lambda(\Lambda)} \\ &= \left[\Lambda \frac{\partial}{\partial \Lambda} + \frac{d\lambda}{d \ln \Lambda} \frac{\partial}{\partial \lambda} + \frac{dm}{d \ln \Lambda} \frac{\partial}{\partial m} \right] \langle O_i \rangle_{\Lambda, m, \lambda}. \end{aligned} \quad (39)$$

カットオフの変更とラグランジアンのパラメータ変更を組み合わせ、指定した長距離の物理を不変に保つ変換が、くりこみ群変換である。式 (39) に現れるパラメータの変化率は、付録 A26 で β 関数として定式化する。

くりこみ群変換で何を不変に保つのかという精密な意味は、付録 A1 で説明する。

ただし a は物理ではないため、物理量を一定に保ちながら最終的に $a \rightarrow 0$ を取る。ここでの極限は、単に格子間隔を小さくするだけではない。格子間隔 a を変えると、紫外カットオフ $\Lambda \simeq 1/a$ も変わり、短距離揺らぎの寄与も変わる。そこで、質量や特定の相関関数など有限個の物理量をくりこみ条件として選び、 $\Lambda = 1/a$ を無限大へ送る間もそれらが一定値を保つように、裸の質量 m と裸の結合定数 λ を Λ に応じて動かす：

$$\Lambda = \frac{1}{a} \longrightarrow \infty, \quad \mathcal{O}_i^{\text{phys}}(m(\Lambda), \lambda(\Lambda); \Lambda) = \mathcal{O}_i^{\text{input}}, \quad i = 1, \dots, N_{\text{input}}. \quad (40)$$

そのため無次元の裸結合定数と裸質量パラメータを $\lambda(a), m(a)$ と書くと、いくつかの物理的な相関関数や質量比を実験値または入力値に合わせながら、

$$a \rightarrow 0, \quad \Lambda = \frac{1}{a} \rightarrow \infty, \quad \lambda(a) \rightarrow \lambda^*, \quad m(a) \rightarrow m^* \quad (41)$$

となるように調整する必要がある。これが、くりこみながら取る量子的な連続極限である。古典的には差分が微分に戻ればよいが、量子論では短距離自由度の効果を吸収するようにパラメータを動かしながら極限を取る点が本質的である。ここで m^* は物理的な粒子質量ではなく、無次元裸質量の固定点または臨界値を表す。物理的な質量を有限に保つには、格子単位の相関長 ξ_{lat} を発散させながら $a\xi_{\text{lat}}$ を一定に保つように、 $m(a)$ の m^* からのずれを調整する。

この調整を有限個のくりこみ条件で済ませられることが、くりこみ可能性の要点である。摂動論では、各次数で現れる紫外発散を有限個の演算子に対応する反項、すなわち有限個の裸パラメータの再定義へ吸収できる理論を、摂動的にくりこみ可能という。カットオフを除くたびに独立な新しい演算子と結合定数を無限個調整しなければならないなら、有限個の入力から理論の予言を定められず、摂動的にはくりこみ不可能である。ただし、そのような理論も有限のカットオフを持つ有効場理論としては、求める精度に応じて有限個の高次元演算子を残すことで系統的に扱える。

2.4 Wilson 型くりこみ群とユニバーサリティ

Wilson 型くりこみ群では、紫外カットオフ Λ 近傍の短距離自由度を積分消去し、座標と場を再スケールして、作用がどのように変化するかを追う。この変換のもとで作用は、場の種類と対称性が許すすべての局所演算子を含む無限次元の理論空間を流れる。固定点の近くでは各演算子は関連、周辺、非関連な方向に分類され、長距離へ進むほど非関連な結合の影響は小さくなる [13]。

したがって、同じ固定点へ流れ込み、有限個の関連な結合を同じ物理量に合わせれば、同じ長距離の物理を与える微視的ラグランジアンや作用は無限に存在する。これがユニバーサリティであり、連続極限の物理は微視的な離散化の細部よりも、時空次元、場の種類、内部対称性やゲージ対称性などによって決まる。格子作用はこの無数の微視的作用の 1 つであり、局所性と必要な対称性を保ち、関連なパラメータを適切に調整すれば、同じ普遍性クラスの連続理論へ近づく。

Symanzik の描像では、4 次元格子理論の長距離での振る舞いを式 (209) の連続有効作用で表す。 $\mathcal{O}_i$ は格子作用の対称性が許す $d_i > 4$ の演算子であり、現象論的な有効場の理論という高次元演算子、Wilson 型くりこみ群という非関連演算子に対応する。異なる格子作用は主として係数 c_i が異なるため、有限 a でのカットオフ効果は異なるが、 $a \rightarrow 0$ では同じ連続理論へ収束する。主要な c_i を打ち消すよう格子作用を選ぶのが Symanzik 改善である [14]。

3 SU(3) 格子ゲージ理論

3.1 $U(1)$ ゲージ対称性と Maxwell–Dirac 理論

ゲージ理論の出発点として、電磁場と荷電 Dirac 場の系を考える。ここでは変換則を見やすくするため Minkowski 時空で書くが、Euclid 時空へ移ってもゲージ変換と共変微分の構造は変わらない。Maxwell–Dirac ラグランジアンは

$$\begin{aligned}\mathcal{L}_{\text{MD}} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi, \\ D_\mu &= \partial_\mu + ieA_\mu, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\end{aligned}\tag{42}$$

である。 α が定数なら、位相回転 $\psi \rightarrow e^{i\alpha}\psi$ は大域的 $U(1)$ 対称性であり、Noether 電流 $j^\mu = \bar{\psi}\gamma^\mu\psi$ の保存に対応する。

次に、位相の基準を時空点ごとに独立に選び、 $\alpha = \alpha(x)$ とする。普通の微分には $\partial_\mu\alpha$ が現れるため、 $\partial_\mu\psi$ だけでは局所変換に対して共変にならない。そこで

$$\begin{aligned}\psi(x) &\longrightarrow e^{i\alpha(x)}\psi(x), & \bar{\psi}(x) &\longrightarrow \bar{\psi}(x)e^{-i\alpha(x)}, \\ A_\mu(x) &\longrightarrow A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x)\end{aligned}\tag{43}$$

と同時に変換する。このとき

$$D_\mu\psi(x) \longrightarrow e^{i\alpha(x)}D_\mu\psi(x), \quad F_{\mu\nu} \longrightarrow F_{\mu\nu}\tag{44}$$

となり、式 (42) は局所 $U(1)$ 変換のもとで不変である。したがって電磁ポテンシャル A_μ は、単に新しいベクトル場を足したのではなく、異なる時空点で選ばれた位相の基準を比較するための接続として現れる。ゲージ変換は物理状態を変える変換というより、同じ物理を異なる局所基底で記述する冗長性である。

3.2 Yang–Mills 理論と内山の一般化

Yang と Mills は 1954 年、電磁気学の局所 $U(1)$ 対称性を、核子のアイソスピンに作用する局所 $SU(2)$ 変換へ一般化した [6]。一般の非可換 Lie 群 G に対し、物質場を

$$\psi(x) \longrightarrow \Omega(x)\psi(x), \quad \Omega(x) \in G \quad (45)$$

と変換する。ゲージ場を Lie 代数に値をもつ行列

$$A_\mu(x) = A_\mu^a(x)T^a, \quad [T^a, T^b] = if^{abc}T^c \quad (46)$$

と表す。 T^a は群の生成子であり、抽象的な記号ではあるが、行列表現を選べば具体的な行列になる。例えば $SU(2)$ の基本表現ではパウリ行列

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (47)$$

を用いて $T^a = \sigma^a/2$ と取れる。このとき $[T^a, T^b] = i\epsilon^{abc}T^c$ 、 $\text{Tr}(T^a T^b) = \delta^{ab}/2$ である。したがって $SU(2)$ ゲージ場 A_μ は、各時空点で 3 個の実係数 A_μ^a を持つ、トレースがゼロのエルミート行列として表される。

QCD のゲージ群 $SU(3)$ の基本表現では、対応する生成子として 8 個のゲルマン行列を用いる。コンパクトに書けば

$$\begin{aligned} \lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda^3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ \lambda^5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda^6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \lambda^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda^8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \quad (48)$$

$T^a = \lambda^a/2$ とすれば、同じ規格化 $\text{Tr}(T^a T^b) = \delta^{ab}/2$ を満たす。 $SU(3)$ では $[T^a, T^b] = if^{abc}T^c$ の構造定数 f^{abc} は ϵ^{abc} より複雑だが、交換子が再び 8 個の生成子の線形結合になる点は同じである。従って QCD のゲージ場は $A_\mu = A_\mu^a \lambda^a/2$ というトレースがゼロの 3×3 エルミート行列であり、8 個の成分 A_μ^a が 8 種類のグルーオンに対応する。生成子の基底には任意性があるが、交換関係とトレース規格化を保つ限り物理的内容は変わらない。

このような行列値のゲージ場に対して、共変微分を

$$D_\mu = \partial_\mu + igA_\mu \quad (49)$$

と定義する。 $D_\mu \psi \rightarrow \Omega D_\mu \psi$ を要求すると、ゲージ場の変換則は

$$A_\mu \longrightarrow \Omega A_\mu \Omega^\dagger + \frac{i}{g}(\partial_\mu \Omega)\Omega^\dagger \quad (50)$$

に決まる。場の強さは共変微分の交換子から

$$\begin{aligned} [D_\mu, D_\nu] &= igF_{\mu\nu}, \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \end{aligned} \quad (51)$$

と得られ、 $F_{\mu\nu} \rightarrow \Omega F_{\mu\nu} \Omega^\dagger$ と変換する。したがって

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2} \text{Tr} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \quad (52)$$

はゲージ不変である。 $U(1)$ では消えていた交換子項が残るため、非可換ゲージ場は物質がなくても自己相互作用し、場の方程式は非線形になる。

内山龍雄は1956年、この構成を特定の群に限らない一般原理として整理した[7]。まず定数パラメータをもつ Lie 群の大域的対称性を考え、そのパラメータを時空の任意関数へ置き換える。局所化した対称性を保つことを要求すると、新しいゲージ場、その変換則、物質場との結合および共変微分が導かれる。内山は同じ枠組みで電磁場、Yang–Mills 場、局所 Lorentz 変換と重力場を比較した。この見方では、ゲージ場とは局所対称性を実現するために任意に付け足す場ではなく、異なる点の内部空間を比較する接続である。

3.3 平行移動子によるゲージ共変差分

連続 Dirac 作用の運動項を格子化すると、異なる格子点にある場の積が必ず現れる。例えば対称差分

$$\partial_\mu \psi(x) \longrightarrow \frac{\psi(x + a\hat{\mu}) - \psi(x - a\hat{\mu})}{2a} \quad (53)$$

を用いると、 $\bar{\psi}(x)\gamma_\mu\partial_\mu\psi(x)$ は $\bar{\psi}(x)\gamma_\mu\psi(x + a\hat{\mu})$ や $\bar{\psi}(x)\gamma_\mu\psi(x - a\hat{\mu})$ を含む。ところが局所ゲージ変換のもとで、例えば前者は

$$\bar{\psi}(x)\gamma_\mu\psi(x + a\hat{\mu}) \longrightarrow \bar{\psi}(x)\Omega^\dagger(x)\gamma_\mu\Omega(x + a\hat{\mu})\psi(x + a\hat{\mu}) \quad (54)$$

と変換する。 $\Omega(x)$ と $\Omega(x + a\hat{\mu})$ は独立なので、素朴な差分は局所ゲージ対称性を保たない。

異なる点の場は、それぞれの点における独立な内部空間の基底で表されている。2点の場を比較するには、経路 C に沿う平行移動子

$$W_C(x, y) = \mathcal{P} \exp \left[ig \int_{C:x \rightarrow y} dz^\mu A_\mu(z) \right] \quad (55)$$

を間に入れる。この平行移動子は Wilson line、または Schwinger line と呼ばれる。 $\mathcal{P}$ は、経路に沿って行列を並べる経路順序積である。非可換理論では異なる点の A_μ が一般に可換でないため、 $\mathcal{P}$ を省略できない。この平行移動子は

$$W_C(x, y) \longrightarrow \Omega(x)W_C(x, y)\Omega^\dagger(y) \quad (56)$$

と変換する。

そこで、式 (53) を

$$(\nabla_\mu^W \psi)(x) = \frac{1}{2a} [W_C(x, x + a\hat{\mu})\psi(x + a\hat{\mu}) - W_C(x, x - a\hat{\mu})\psi(x - a\hat{\mu})] \quad (57)$$

と置き換える。式 (56) から $\nabla_\mu^W \psi(x) \rightarrow \Omega(x)\nabla_\mu^W \psi(x)$ となるため、 $\bar{\psi}(x)\gamma_\mu\nabla_\mu^W \psi(x)$ はゲージ不変である。さらに短い直線経路について Wilson line を a で展開すると

$$\nabla_\mu^W \psi(x) = (\partial_\mu + igA_\mu(x))\psi(x) + O(a^2) = D_\mu\psi(x) + O(a^2). \quad (58)$$

したがって式 (57) は、連続理論の minimal coupling を局所ゲージ対称性を保ったまま差分化したものである。

非可換理論では、経路順序積を含むこの平行移動子を Wilson line と呼ぶ。閉路に沿ってトレースを取れば Wilson loop になり、直線経路で $y = x + a\hat{\mu}$ とすれば、格子上の1本のリンク変数になる。point splitting 法と一般の複合演算子への応用は付録 A29 で説明する。

これで物質場の運動項は、局所ゲージ対称性を保った差分として書ける。しかし Yang–Mills 作用には、物質場との結合とは別に、ゲージ場自身の運動項 $\text{Tr} F_{\mu\nu} F^{\mu\nu}$ がある。微分を直接使わず、リンク変数だけからこの項に対応する局所ゲージ不変な作用を構成することが次の課題である。

3.4 ゲージ対称性を保った格子ゲージ作用

前節では Wilson line を用いて物質場の運動項を共変に差分化した。次に、ゲージ場自身の運動項を格子上で構成する。ゲージ理論を有限自由度の積分として非摂動的に定義し、数値計算できる形にするには時空を離散化したいが、局所ゲージ対称性は素朴な離散化とは相性が悪い。連続ゲージ場 $A_\mu(x)$ は、式 (50) のように微分項を含んで変換する接続であり、通常の局所演算子のように 1 点だけで共変変換する場ではない。 A_μ を格子点に置き、そこから $F_{\mu\nu}$ を単純な差分で作るだけでは、有限の格子間隔で局所ゲージ対称性が自動的に保たれるわけではない。

一方、ユニバーサリティの考え方によれば、局所性と対象とする連続理論の対称性を保ち、関連なパラメータを適切に調整すれば、格子作用の微視的な細部が異なっても、連続極限では同じ長距離の物理に到達すると期待される。したがって、格子ゲージ作用には少なくとも 2 つの条件を課したい。第 1 に、有限の格子間隔 a でも局所ゲージ対称性を厳密に保つこと、第 2 に、 $a \rightarrow 0$ で目的とする連続 Yang–Mills 作用へ戻ることである。この 2 つの条件を同時に満たす離散化を構成する。

そこで式 (55) の最短の直線経路に対応するリンク変数

$$U_\mu(n) \in SU(3), \quad U_\mu(n) \simeq e^{ia g A_\mu(x+a\hat{\mu}/2)} \quad (59)$$

を天下一りに導入する。これが格子時空でのゲージ場の役割を果たす。変換則を

$$U_\mu(n) \longrightarrow \Omega(n) U_\mu(n) \Omega^\dagger(n + \hat{\mu}) \quad (60)$$

とすれば

$$U_\mu(n) \psi(n + \hat{\mu}) \longrightarrow \Omega(n) U_\mu(n) \psi(n + \hat{\mu}) \quad (61)$$

となり、隣の点の場を n へ運んで比較できる。

格子ゲージ理論では、ゲージ場は点ではなく辺に置く。

3.5 最小閉路としてのプラケット

前小節でリンク変数によりゲージ対称性を保った離散化の方針を得た。ゲージ場の運動項を作るには、リンクからゲージ不変な局所量を構成する必要がある。ゲージ不変な局所量を作るには、リンクを閉じた経路に沿って掛ければよい。最小の正方形を回るプラケットを

$$U_{\mu\nu}(n) = U_\mu(n) U_\nu(n + \hat{\mu}) U_\mu^\dagger(n + \hat{\nu}) U_\nu^\dagger(n) \quad (62)$$

と定義する。ゲージ変換すると

$$U_{\mu\nu}(n) \longrightarrow \Omega(n) U_{\mu\nu}(n) \Omega^\dagger(n) \quad (63)$$

なので、トレース $\text{Re Tr } U_{\mu\nu}$ はゲージ不変である。

3.6 Wilson 作用

前小節で作ったゲージ不変なプラケットを格子全体で足し合わせ、ゲージ場の作用を定義する。純粋ゲージ理論の Wilson 作用は

$$S_g[U] = \beta \sum_n \sum_{\mu < \nu} \left(1 - \frac{1}{N_c} \text{Re Tr } U_{\mu\nu}(n) \right), \quad N_c = 3. \quad (64)$$

QCD では慣習的に $\beta = 2N_c/g_0^2 = 6/g_0^2$ である。ここで β は逆温度ではなく裸のゲージ結合を表す記号である。この作用は Wilson による格子ゲージ理論の定式化 [8] の基本である。

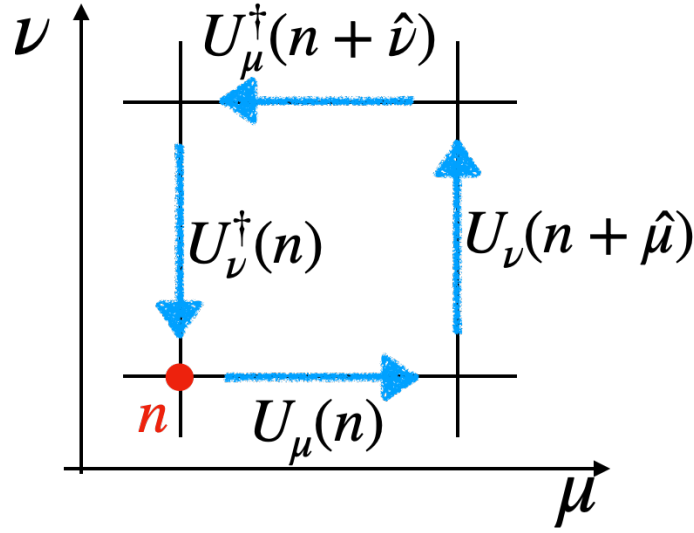


図 4: 格子上的最小閉路（プラケット）。矢印の向きに進むリンクは U 、逆向きは $U^\dagger$ で表す。閉路積のトレースを取るとゲージ不変量になる。

3.7 プラケットと連続ゲージ作用

記法を簡単にするためリンクを始点 x で

$$U_\mu(x) = e^{ig_0 A_\mu(x)} \quad (65)$$

と書く。プラケットは

$$U_{\mu\nu}(x) = e^{ig_0 A_\mu(x)} e^{ig_0 A_\nu(x+a\hat{\mu})} e^{-ig_0 A_\mu(x+a\hat{\nu})} e^{-ig_0 A_\nu(x)}. \quad (66)$$

ずれた点のゲージ場を

$$A_\nu(x+a\hat{\mu}) = A_\nu(x) + a\partial_\mu A_\nu(x) + O(a^2), \quad (67)$$

$$A_\mu(x+a\hat{\nu}) = A_\mu(x) + a\partial_\nu A_\mu(x) + O(a^2) \quad (68)$$

と Taylor 展開する。指数行列は一般に可換ではないため、Baker–Campbell–Hausdorff 公式

$$e^X e^Y = \exp \left[X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \frac{1}{12}[Y, [Y, X]] + \cdots \right] \quad (69)$$

を使う。式 (66) では $O(a)$ 項が互いに消え、Taylor 展開から微分項、非可換な積から交換子項が残る。合わせると

$$U_{\mu\nu}(n) = \exp [ig_0 a^2 F_{\mu\nu}(x) + O(a^3)], \quad (70)$$

ただし

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig_0 [A_\mu, A_\nu]. \quad (71)$$

ここで交換子項の符号は共変微分 $D_\mu = \partial_\mu + ig_0 A_\mu$ の規約に対応する。 $SU(N_c)$ 生成子はトレースレスなので $\text{Tr} F_{\mu\nu} = 0$ であり、指数関数をさらに展開すると

$$\text{Re Tr } U_{\mu\nu} = \text{Re Tr} \left[1 + ig_0 a^2 F_{\mu\nu} - \frac{g_0^2 a^4}{2} F_{\mu\nu}^2 + O(a^6) \right] \quad (72)$$

$$= N_c - \frac{g_0^2 a^4}{2} \text{Tr} F_{\mu\nu}^2 + O(a^6). \quad (73)$$

この展開から

$$1 - \frac{1}{N_c} \text{Re Tr } U_{\mu\nu} = \frac{g_0^2 a^4}{2N_c} \text{Tr } F_{\mu\nu}^2 + O(a^6) \quad (74)$$

を得る。これを Wilson 作用 (64) へ代入し、格子和 $a^4 \sum_n \rightarrow \int d^4x$ を用いると

$$S_g \xrightarrow{a \rightarrow 0} \frac{\beta g_0^2}{2N_c} \int d^4x \sum_{\mu < \nu} \text{Tr } F_{\mu\nu}^2. \quad (75)$$

一方、Euclid 時空の Yang–Mills 作用は

$$S_{\text{YM}} = \int d^4x \sum_{\mu < \nu} \text{Tr } F_{\mu\nu}^2 = \frac{1}{2} \int d^4x \sum_{\mu, \nu} \text{Tr } F_{\mu\nu}^2 \quad (76)$$

である。両者を一致させる条件は

$$\beta = \frac{2N_c}{g_0^2}, \quad \text{SU}(3) : \quad \beta = \frac{6}{g_0^2} \quad (77)$$

である。重要なのは、連続作用を再現するだけでなく、有限 a でもゲージ不変性が厳密に保たれていることである。

リンク変数を用いる理由

A_μ を単純に差分化する代わりにリンク変数を使うのは、離散化誤差が残る有限 a でも、局所ゲージ対称性を厳密に保つためである。

4 格子フェルミオン

4.1 ナイーブ差分とダブリング

ゲージ場については、リンク変数を使えば有限の格子間隔でも局所ゲージ対称性を保てた。次は、その背景上を伝播するクォークを格子に置く。連続理論の Dirac 演算子 $\gamma_\mu D_\mu + m$ を有限次元の行列 $D[U]$ へ置き換え、局所性とゲージ共変性を保ちながら、 $a \rightarrow 0$ で正しいフェルミオンのスペクトルを再現することが目標である。ここで、連続場の理論における D_μ は方向添字 μ を持つ共変微分と呼ばれる。一方、格子上の有限次元行列 $D[U]$ は、スピノル、色および格子点の添字に作用する Dirac 演算子と呼ばれる。同じ D を用いるが、 D_μ と $D[U]$ は区別しておく。

ダブリングの原因はゲージ相互作用を入れる前から存在するので、まず $U_\mu(n) = 1$ の自由場を調べる。並進対称な格子では平面波が Dirac 演算子の固有関数となる。ただし格子運動量は連続理論のように無限に広がるのではなく、第 1 ブリルアンゾーン $-\pi/a < p_\mu \leq \pi/a$ に収まり、端どうしが周期的に同一視される。質量ゼロでは Dirac 演算子の零点が低エネルギーのフェルミオンに対応するため、格子上で零点がいくつ現れるかを数えることが、粒子種を数えることになる。

連続理論では質量ゼロの逆伝播関数 $i\gamma_\mu p_\mu$ の零点は $p_\mu = 0$ だけである。そこで格子差分にも、 $p_\mu \simeq 0$ で ip_μ を再現し、余分な零点を持たないことを期待する。最も自然に見えるのは、前後の格子点を対等に扱う対称差分である。これは局所的で、方向の偏りがなく、古典的な離散化誤差も $O(a^2)$ である。しかし、この素朴な選択がブリルアンゾーンの端に余分な零点を作る。以下でその過程を確認する。

連続ユークリッド作用を

$$S_f = \int d^4x \bar{\psi}(\gamma_\mu D_\mu + m)\psi \quad (78)$$

とする。自由場では $D_\mu = \partial_\mu$ である。ゲージ場を戻すときには、小節 3.3 で導入した平行移動子と同様に、隣接点のフェルミオン場をリンク変数で平行移動して結ばばよい。まず自由場で対称差分

$$\partial_\mu \psi(x) \longrightarrow \frac{\psi_{n+\hat{\mu}} - \psi_{n-\hat{\mu}}}{2a} \quad (79)$$

を使う。平面波 $\psi_n = u(p)e^{iap \cdot n}$ を代入すると

$$\frac{e^{iap_\mu} - e^{-iap_\mu}}{2a} = \frac{i}{a} \sin(ap_\mu) \quad (80)$$

だから、運動量空間の Dirac 演算子は

$$D_{\text{naive}}(p) = m + \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(ap_{\mu}). \quad (81)$$

$m = 0$ で零点を探すと、各方向で $ap_{\mu} = 0$ だけでなく π でも $\sin(ap_{\mu}) = 0$ になる。4 次元では $2^4 = 16$ 個の零点があり、1 つのフェルミオンを入れたつもりが 16 種現れる。これが fermion doubling である。ダブラーが連続極限でも消えない理由は、付録 A28 で詳しく説明する。

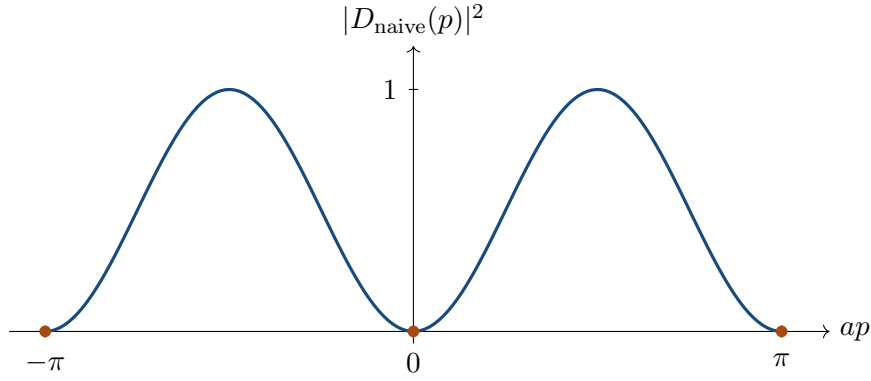


図 5: 1 次元のナイーブ Dirac 演算子の零点。 $-\pi/a$ と $+\pi/a$ は周期的に同一の境界点なので、独立な零点は中央と境界の 2 つである。4 次元では各方向の組合せにより 2^4 個になる。

4.2 Nielsen–Ninomiya 定理

自由な質量ゼロフェルミオンに対する標準的な Euclid 格子 Dirac 演算子 D を考える。この定理で同時に課す仮定を、ここでは次のように明示する。

1. **並進対称性:** $D_{nm} = D(n - m)$ であり、運動量空間で $D(p)$ と書ける。
2. **局所性:** $D(n - m)$ は距離とともに十分速く減衰する。このため $D(p)$ はブリルアンゾーン上の滑らかで周期的な関数になる。
3. **γ_5 -Hermiticity:** $D^\dagger = \gamma_5 D \gamma_5$ を満たす。
4. **通常のカイラル対称性:** 有限の格子間隔でも $\{\gamma_5, D\} = 0$ を満たす。
5. **正しい連続極限:** $p \simeq 0$ で $D(p) = i\gamma_\mu p_\mu + O(ap^2)$ となり、目的とする質量ゼロ Dirac フェルミオンを再現する。

これらを満たす $D(p)$ は、ブリルアンゾーン内に原点以外の零点を持たざるを得ない。従って、上の仮定と「ダブラーが存在しないこと」は両立せず、単独のカイラルフェルミオンを格子上に置けない。したがって格子フェルミオンは、何を厳密に保ち、何を有限 a で犠牲にするかの選択になる。これが Nielsen–Ninomiya の禁止定理 [16] の内容である。

定理の背景は、格子運動量が周期的なブリルアンゾーン上にあることにある。質量ゼロの Dirac 演算子の零点を 1 つだけ残そうとしても、局所的で滑らかな運動量表示を保つ限り、反対のカイラリティーを持つ零点が別の場所に現れる。零点に付随するカイラリティーの総和がブリルアンゾーン全体で相殺されるためである。従って、ナイーブな離散化で現れるダブラーは単なる差分公式の選び方の失敗ではなく、要求した性質の組合せから生じる。

ここで禁じられているのは、有限の格子間隔で、連続理論と同じ形のカイラル対称性を保ち、局所かつダブラーのない Dirac 演算子を同時に実現することである。どれかの仮定を緩めれば格子フェルミオンは構成できる。

4.3 Wilson フェルミオン

Wilson フェルミオンでは、ナイーブ Dirac 演算子に、位置空間の 2 階差分に対応する Wilson 項

$$\frac{r}{a} \sum_{\mu} (1 - \cos(ap_{\mu})) \quad (82)$$

を加える。ここで r は無次元の Wilson パラメータであり、ダブラーに与える質量の大きさを調整する。 $r = 0$ ではナイーブフェルミオンに戻るためダブラーは除去されない。通常は $r = 1$ を選ぶ。 $r \neq 0$ なら、 $p_{\mu} \simeq 0$ では Wilson 項は $O(ap^2)$ なので連続極限で消える一方、 $ap_{\mu} = \pi$ のダブラーには各方向につき $2r/a$ の質量を与える。代償として、有限 a ではカイラル対称性を明示的に破る。

歴史的には、Wilson は 1974 年の論文「Confinement of quarks」[8] の段階では、少なくとも論文にダブリング問題へ気づいた記述を残していない。同論文のフェルミオン作用は対称 1 階差分のままであり、2 階差分の Wilson 項も書かれていない。Wilson 項が導入されたのは、翌 1975 年に行われた Erice での講義録「Quarks and Strings on a Lattice」[9] である。従って、1974 年の格子ゲージ理論の定式化と、現在 Wilson フェルミオンと呼ばれる処方は、同じ文献で同時に完成したものではない。

もう少し具体的に、Wilson-Dirac 演算子を自由場で

$$D_W(p) = m + \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(ap_{\mu}) + \frac{r}{a} \sum_{\mu} [1 - \cos(ap_{\mu})] \quad (83)$$

と書く。 $ap_{\mu} \simeq 0$ では

$$\frac{r}{a} [1 - \cos(ap_{\mu})] = \frac{ra}{2} p_{\mu}^2 + O(a^3 p_{\mu}^4) \quad (84)$$

なので物理モードへの影響は $a \rightarrow 0$ で消える。一方、 n 方向の運動量成分が $ap_{\mu} = \pi$ にあるダブラーには

$$m_{\text{doubler}} = m + \frac{2rn}{a} \quad (85)$$

という質量が加わり、連続極限で無限に重くなって物理的な自由度から分離する。有限 a でカイラル対称性を破るため、裸質量には加法的くりこみが生じる。

Nielsen-Ninomiya 定理のため、格子フェルミオン作用は、ダブラーの除去、カイラル対称性、局所性、計算量などを同時に理想的な形で満たすことができず、必ず一長一短を持つ。代表的な定式化は、この制約への異なる応答として整理できる。

Wilson フェルミオン ダブラーを $O(1/a)$ の質量で取り除く代わりに、有限 a で通常のカイラル対称性を破る。

スタッガードフェルミオン スピン成分を格子点へ分散してダブラーを減らし、カイラル対称性の一部を保つ。ここで残る taste (テイスト) は、連続場の理論におけるフレーバーに相当する内部自由度である。4 次元では 1 つのスタッガード場から 4 つのテイスト、すなわち似た質量を持つ 4 つのフレーバーに相当する自由度が現れる。行列の自由度が少ないため計算は比較的軽いが、有限の格子間隔ではテイスト対称性が破れて 4 つの質量が完全には一致しない。また、1 または 2 フレーバーを表すためにフェルミオン行列式の 4 乗根を用いる rooting と、その有限格子間隔での局所性・連続極限の扱いには注意が必要である。

ドメインウォールフェルミオン 余分な次元を導入し、その境界に低エネルギーのカイラルフェルミオンを実現する。カイラル対称性を高い精度で保てる一方、5 次元の自由度を扱うため Wilson 型より計算量が大きい。

オーバーラップフェルミオン くりこみ群の完全作用に由来する Ginsparg-Wilson 関係を満たし、格子上で修正された厳密なカイラル対称性を実現する特別な作用である。ただし、Dirac 演算子に含まれる行列符号関数の評価には多数回の線形方程式の求解が必要となり、数値計算コストは Wilson 型やスタッガード型に比べて劇的に高い。

したがって最良の離散化が 1 つに決まるわけではなく、対称性、離散化誤差、計算量のどれを重視するかによって選択が変わる。

5 第1講のまとめ

本文では、Euclid 経路積分から出発し、スカラー場、ゲージ場、フェルミオンを格子上で定義するところまでを見た。動的クォークを含む QCD で現れるフェルミオン行列式と擬フェルミオン表示は、それぞれ付録 A8 と付録 A9 にまとめる。ここまでの要点は次の 4 つである。

第1講の要点

1. Wick 回転後の経路積分は、条件がよければ確率分布と同じ形になる。
2. 格子は自由度を有限にし、局所対称性を保った非摂動的定義を与える。
3. 連続極限は単なる差分の極限ではなく、物理量を固定した臨界極限である。
4. 格子フェルミオンでは、ダブラー、カイラル対称性、局所性、計算量の間に一長一短がある。

残る問題は、格子作用が定める高次元分布から代表的な配位をどう効率よく作るかである。

第2講

格子場理論におけるサンプリングと機械学習

6 モンテカルロ積分とマルコフ連鎖

6.1 台形則・区分求積と次元の呪い

1変数の積分なら、積分区間を細かく分ける区分求積法が自然である。 $[a, b]$ を M 等分し、 $h = (b-a)/M$ とすると、台形則は

$$\int_a^b dx f(x) \approx h \left[\frac{1}{2}f(a) + \sum_{k=1}^{M-1} f(a+kh) + \frac{1}{2}f(b) \right] \quad (86)$$

である。 f が2回連続微分可能なら、誤差は

$$|I - I_M| \leq \frac{(b-a)^3}{12M^2} \max_{x \in [a,b]} |f''(x)| \quad (87)$$

と評価できる。1次元では、分割数を増やせば積分値へ系統的に近づける。

しかし、 D 変数の積分で各変数を M 分割し、同じ規則を直積として用いると、評価点数は

$$N_{\text{grid}} = M^D \quad (88)$$

になる。各方向の離散化誤差を $O(M^{-2})$ まで小さくする場合、固定した D に対して $M \sim \epsilon^{-1/2}$ が必要なので、計算量は概ね $N_{\text{grid}} \sim \epsilon^{-D/2}$ と増える。これが区分求積における次元の呪いである。実スカラー場では積分領域が非有界なので、実際には各 ϕ_n の範囲も有限に切る必要があり、格子幅による誤差に加えて切断誤差も管理しなければならない。

格子場理論では D は時空次元ではなく、積分する場の自由度数である。例えば 16^4 格子の実スカラー場でも $D = 65,536$ 、 32^4 格子なら $D = 1,048,576$ である。各変数をわずか10点で評価しても、必要な評価点数はそれぞれ $10^{65,536}$ 点、 $10^{1,048,576}$ 点になる。

所要時間の目安を見積もるため、富岳の倍精度理論ピーク性能を丸めて 5×10^{17} 演算/秒とし、さらに1評価点をわずか1演算で処理できるという、現実にはあり得ないほど楽観的な仮定を置く²。それでも、わずか $D = 100$ に対して

$$T_{D=100} \gtrsim \frac{10^{100}}{5 \times 10^{17}} \text{ s} \simeq 2.0 \times 10^{82} \text{ s} \quad (89)$$

となる。これは約 6.3×10^{74} 年であり、宇宙年齢を約 1.38×10^{10} 年とすれば、その約 4.6×10^{64} 倍である。 16^4 格子なら同じ理想化のもとでも約 $10^{65,501}$ 宇宙年齢が必要であり、実際の1点あたりの作用評価には多数の演算が要するため、真の所要時間はさらに長い。疎格子法や準 Monte Carlo 法は低い有効次元や特別な滑らかさを利用できる場合に有効だが、一般の相互作用する格子場の積分を直積型の区分求積で計算することは現実的でない。

Monte Carlo 積分では、全格子点を走査する代わりに、重要な領域から N 個の配位を抽出する。独立標本に対する統計誤差 $O(N^{-1/2})$ の収束指数は D に直接依存しない。もちろん誤差の係数や標本生成の難しさは自由度数とともに悪化するが、 M^D 個の格子点を列挙しないことが、場の理論で Monte Carlo 法を使う根本的な理由である。

²理化学研究所計算科学研究センターによる富岳の倍精度理論ピーク性能は、通常モードで 488 PFLOPS、ブーストモードで 537 PFLOPS である。<https://www.r-ccs.riken.jp/en/fugaku/about/>

6.2 独立標本による期待値の推定

格子場理論で確率変数となるのは、1つの格子点の場合だけではなく、全格子点の場をまとめた配位

$$\varphi = (\varphi_n)_{n \in \Lambda} \in \mathbb{R}^V \quad (90)$$

である。有限格子 Λ 上の実スカラー場では、分配関数と観測量 $O[\varphi]$ の期待値は

$$\begin{aligned} Z_{\text{lat}} &= \int \prod_{n \in \Lambda} d\varphi_n e^{-S_{\text{lat}}[\varphi]}, \\ \langle O \rangle &= \frac{1}{Z_{\text{lat}}} \int \prod_{n \in \Lambda} d\varphi_n O[\varphi] e^{-S_{\text{lat}}[\varphi]} \end{aligned} \quad (91)$$

と書ける。評価したい O は、場の1点関数、2点相関関数、作用密度、磁化など、全配位から計算される関数である。連続理論の形式的な汎関数積分は、格子化によって式 (91) の V 変数の通常の積分になる。

Ising 模型では確率変数は全スピン配置 $S = (s_1, \dots, s_N) \in \{\pm 1\}^N$ であり、積分の代わりに全配置の和を取る。

$$\begin{aligned} Z_{\text{Ising}} &= \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-\beta \mathcal{H}_{\text{Ising}}(\{s\})}, \\ \langle O \rangle &= \frac{1}{Z_{\text{Ising}}} \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} O(\{s\}) e^{-\beta \mathcal{H}_{\text{Ising}}(\{s\})}. \end{aligned} \quad (92)$$

例えば $O = N^{-1} \sum_i s_i$ なら磁化を、 $O = s_i s_j$ ならスピン相関を評価する。

場やスピンを点ごとに独立に生成できない原因は、作用またはハミルトニアンにある隣接点間の結合である。式 (31) の運動項を展開すると、 $-\sum_{n,\mu} \varphi_n \varphi_{n+\hat{\mu}}$ というホッピング項が現れる。他の点を固定したとき、1点 n の条件付き分布は

$$\begin{aligned} p(\varphi_n \mid \{\varphi_m\}_{m \neq n}) &\propto \exp[-S_n(\varphi_n; \varphi_{\partial n})], \\ S_n &= \left(d + \frac{(am_0)^2}{2} \right) \varphi_n^2 + \frac{\lambda_0}{4!} \varphi_n^4 - \varphi_n \sum_{\mu=1}^d (\varphi_{n+\hat{\mu}} + \varphi_{n-\hat{\mu}}) \end{aligned} \quad (93)$$

となり、近傍の場に依存する。1次元 Ising 模型でも

$$p(s_i \mid \{s_j\}_{j \neq i}) \propto \exp[\beta s_i \{J(s_{i-1} + s_{i+1}) + h\}] \quad (94)$$

であり、隣のスピンを知らなければ s_i の分布は決まらない。この依存関係が全格子へ連なり、相関を持つ巨大次元分布からのサンプリングを難しくする。

ホッピング項を取り除けば、重みは格子点ごとの因子の積に分解し、各点を独立に生成できる。しかしその場合、異なる点の連結相関は消え、励起が格子を伝播する機構も失われる。したがって、サンプリングを難しくしている点間結合は、同時に場の伝播と物理的相関を生む項でもある。

一般に、規格化定数 Z を知らなくても、 $\pi(x) = e^{-S(x)}/Z$ から独立な全配位 x_i が得られれば

$$\langle O \rangle_\pi = \int dx \pi(x) O(x) \approx \frac{1}{N} \sum_{i=1}^N O(x_i) \quad (95)$$

で期待値を計算できる。独立なら誤差は概ね $N^{-1/2}$ で減る。より具体的には、 O の分散が有限なら中心極限定理により

$$\text{Var}(\bar{O}) = \frac{\text{Var}_\pi(O)}{N} \quad (96)$$

となる。したがって、誤差を半分にするには標本数を4倍にする必要がある。

ここで、ボルツマン重みの形に注意する。基準となる配位 x_* より作用が $\Delta S = S(x) - S(x_*) > 0$ だけ大きい配位 x の相対確率は

$$\frac{\pi(x)}{\pi(x_*)} = \exp[-\Delta S] \quad (97)$$

である。作用差が体積に比例して $\Delta S = V \delta s$ と増える場合、 $V = 10^3$ 、 $\delta s = 0.1$ だけでも相対重みは $e^{-100} \simeq 3.7 \times 10^{-44}$ になる。この指数抑制のため、一様な基準測度から見れば、高次元配位空間のほとんどの場所でボルツマン重みは事実上ゼロであり、形式的な配位和や積分領域の大部分は期待値へほとんど寄与しない。

ただし、最小作用の配位だけが支配するわけではない。個々の重みが小さくても配位数が多ければ、その集合の寄与は大きい。分配関数を支配するのは、作用による抑制と配位数（エントロピー）が釣り合う**典型集合**である。典型集合を重点的に調べるのが重要度サンプリングの出発点になるが、一般の π では典型集合から独立な全配位を直接生成することが難しい。

6.3 2点相関関数と質量の測定

配位 $\varphi^{(i)}$ が N 個あれば、観測量は

$$\langle O \rangle \approx \overline{O} = \frac{1}{N} \sum_{i=1}^N O[\varphi^{(i)}] \quad (98)$$

で評価する。例えば零運動量射影した場

$$\Phi(t) = \sum_{\mathbf{x}} \varphi(\mathbf{x}, t) \quad (99)$$

を用いる。Euclid 虚時間での演算子の発展は

$$\Phi(t) = e^{Ht} \Phi(0) e^{-Ht} \quad (100)$$

である。真空を $|0\rangle$ とし、エネルギー固有状態の完全系を挿入すると

$$\begin{aligned} C(t) &= \langle 0 | \Phi(t) \Phi(0) | 0 \rangle \\ &= \sum_n |\langle 0 | \Phi(0) | n \rangle|^2 e^{-(E_n - E_0)t} \end{aligned} \quad (101)$$

となる。虚時間発展では高いエネルギーを持つ状態ほど速く指数関数的に減衰する。以下では $\langle 0 | \Phi | 0 \rangle = 0$ を仮定する。³ Φ と同じ量子数を持つ零運動量状態のうち、最も低いものを $|1\rangle$ とすると、十分大きな t では

$$C(t) \xrightarrow{t \rightarrow \infty} |\langle 0 | \Phi(0) | 1 \rangle|^2 e^{-(E_1 - E_0)t} \equiv A e^{-mt}, \quad m = E_1 - E_0 \quad (102)$$

となる。従って、相関関数の虚時間方向の指数減衰率から質量 m を測定できる。有限の時間方向を持つ格子では、時間境界を回る逆向きの伝播も加わるため、式 (341) の双曲線余弦（cosh）型に適合させ、有効質量が一定になる時間領域と適合範囲への依存性を調べる。

標本平均だけでは統計誤差は分らない。配位がマルコフ連鎖から得られた場合には、自己相関より十分長い幅でデータをまとめた上でジャックナイフ法やブートストラップ法を用いる。サンプル生成、観測量の構成、統計誤差の評価は別の段階であるが、最終結果には3つすべてが必要である。

³自発的対称性の破れなどにより $\langle \Phi \rangle \neq 0$ となる場合、 $n=0$ の項 $|\langle 0 | \Phi | 0 \rangle|^2$ は時間に依存しない。この寄与を除くため、 $C_{\text{conn}}(t) = \langle \Phi(t) \Phi(0) \rangle - \langle \Phi(t) \rangle \langle \Phi(0) \rangle$ を用いる。以下の $C(t)$ を $C_{\text{conn}}(t)$ に読み替えれば、同じ議論が成り立つ。

6.4 重要度サンプリングとマルコフ連鎖 Monte Carlo 法

高次元積分では、積分領域を一様に調べるのではなく、被積分関数への寄与が大きい $\pi(x) \propto e^{-S(x)}$ の領域を重点的に標本化する。この考え方を重要度サンプリングという。しかし、扱いやすい別の分布から全配位を独立に生成する方法では、その分布と目標分布の典型的な配位が重ならないと、ごく少数の標本だけが大きな重みを持つ。格子体積が大きいほどこのオーバーラップ問題は深刻になり、独立標本を直接生成することは難しくなる。

マルコフ連鎖 Monte Carlo 法 (MCMC) では、 π を定常分布として保つ遷移確率を設計し、現在の配位から次の配位を順に作る。標本は相関するが、確率の比を使うため規格化定数 Z は不要である。運動項によって隣接格子点が結合していても、局所的な作用差を用いる更新や、全自由度をまとめて動かす HMC 法を構成できる。独立標本による再重み付けとの違いと、運動項がある場合の詳しい説明は付録 A10 にまとめる。次の小節では、有限状態の模型を用いて、遷移の反復が定常分布へ収束するとはどういうことかを見る。

6.5 Glasgow の天気模型と確率過程の収束

マルコフ連鎖の一般論へ進む前に、有限状態の例を 1 つ計算する。Glasgow のある日の天気を、乾燥 (D) と雨 (W) の 2 状態だけで表す模式模型を考える。ここで用いる数値は実測値ではなく、雨の日が続くやすい状況を表すための単純な例である。明日の天気は今日の天気だけに依存し、遷移確率を

$$\begin{aligned} \Pr(X_{t+1} = D \mid X_t = D) &= \frac{3}{5}, & \Pr(X_{t+1} = W \mid X_t = D) &= \frac{2}{5}, \\ \Pr(X_{t+1} = D \mid X_t = W) &= \frac{1}{5}, & \Pr(X_{t+1} = W \mid X_t = W) &= \frac{4}{5} \end{aligned} \quad (103)$$

とする。

以下では確率分布を縦ベクトルで表す。状態の順序を (D, W) として

$$\mathbf{p}_t = \begin{pmatrix} \Pr(X_t = D) \\ \Pr(X_t = W) \end{pmatrix}, \quad P = \begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix} \quad (104)$$

と定義する。この規約では $P_{ij} = \Pr(X_{t+1} = i \mid X_t = j)$ なので、各列の和が 1 になる。確率分布は

$$\mathbf{p}_{t+1} = P\mathbf{p}_t, \quad \mathbf{p}_t = P^t \mathbf{p}_0 \quad (105)$$

と時間発展する。

長時間後の分布を求めるには、時間発展で変わらない縦ベクトル $\boldsymbol{\pi}$ を

$$P\boldsymbol{\pi} = \boldsymbol{\pi}, \quad \mathbf{1}^\top \boldsymbol{\pi} = 1 \quad (106)$$

から求めればよい。この模型では

$$\boldsymbol{\pi} = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix} \quad (107)$$

である。 P のもう 1 つの固有値は $2/5$ で、対応する固有ベクトルは $(1, -1)^\top$ である。従って任意の初期分布は

$$\mathbf{p}_t = \boldsymbol{\pi} + c \begin{pmatrix} 2 \\ 5 \end{pmatrix}^t \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xrightarrow{t \rightarrow \infty} \boldsymbol{\pi} \quad (108)$$

と定常分布へ収束する。例えば初日が必ず乾燥なら $\mathbf{p}_0 = (1, 0)^\top$ なので $c = 2/3$ である。ここで収束するのは確率分布であり、1 本の時系列の天気が最後には変化しなくなるという意味ではない。定常状態でも D と W の間の遷移は続く。

観測量 O が乾燥日に O_D 、雨の日に O_W を取るなら、 $\mathbf{o} = (O_D, O_W)^T$ として期待値は

$$\mathbb{E}[O(X_t)] = \mathbf{o}^T \mathbf{p}_t \xrightarrow[t \rightarrow \infty]{} \mathbf{o}^T \boldsymbol{\pi} \quad (109)$$

となる。雨なら 1、乾燥なら 0 を返す指示関数では $\mathbf{o} = (0, 1)^T$ であり、雨である確率の期待値は $2/3$ へ近づく。

さらにこの連鎖は、どの状態からも他方へ移れ、自己遷移も持つ。そのため 1 本の十分長い時系列についても

$$\frac{1}{N} \sum_{t=1}^N O(X_t) \xrightarrow[N \rightarrow \infty]{\text{ほとんど確実に}} \mathbf{o}^T \boldsymbol{\pi} \quad (110)$$

が成り立つ。これは、目標分布から独立標本を直接作れなくても、正しく設計したマルコフ連鎖を長く走らせれば期待値を標本平均で求められる、という MCMC の基本原理である。定常分布、詳細釣り合い、エルゴード性の一般的な関係は付録 A30 にまとめる。

6.6 Metropolis–Hastings 法

前節の天気模型では、遷移行列 P を先に与え、その反復によって得られる定常分布を求めた。格子場理論では順序が逆である。目標とする分布 $\pi(x) \propto e^{-S(x)}$ は作用から分かっているが、この分布を定常分布として保ちながら配位空間を探索する遷移則を構成しなければならない。Metropolis–Hastings 法は、現在の配位 x から提案分布 $q(x'|x)$ に従って候補 x' を作り、適切な確率で受容または棄却することで、提案分布の偏りを補正し、 π を定常分布に持つマルコフ連鎖を構成する方法である。

Metropolis 法 Metropolis らは 1953 年、規格化されていない Boltzmann 重みだけを用いて配置空間を探索する方法を示した [3]。元の Metropolis 法では対称な提案 $q(x'|x) = q(x|x')$ を仮定する。現在の配位 x から候補 x' を生成し、作用差 $\Delta S = S(x') - S(x)$ を計算する。候補を

$$A(x \rightarrow x') = \min(1, e^{-\Delta S}) \quad (111)$$

の確率で受容し、棄却した場合は x にとどまる。 $\Delta S \leq 0$ なら必ず受容し、 $\Delta S > 0$ でも確率 $e^{-\Delta S}$ で受容するため、作用を下げる方向だけに動いて局所極小へ閉じ込められることを避けられる。また、目標分布の規格化定数は比で消えるので、 Z を知る必要はない。提案が小さすぎれば高い確率で受容されても移動距離が短く、大きすぎれば棄却が増える。この提案幅の調整が Metropolis 法の基本である。

Metropolis–Hastings 法 Hastings は 1970 年、Metropolis 法を非対称な提案にも適用できる受容率へ一般化した [4]。現在 Metropolis–Hastings 法と呼ぶ形はこの一般化である。提案分布 $q(x'|x)$ から候補 x' を作り、確率

$$A(x \rightarrow x') = \min \left[1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)} \right] \quad (112)$$

で受容する。規格化定数 Z は比で消える。提案が対称 $q(x'|x) = q(x|x')$ なら式 (111) の元の Metropolis 法へ戻る。この受容率が詳細釣り合いを満たすことの導出は、付録 A31 に示す。従って、受容と棄却を合わせた遷移核は π を定常分布に持つ。さらに連鎖がエルゴード的であれば、十分長い連鎖 $X_1, \dots, X_N$ について

$$\frac{1}{N} \sum_{t=1}^N O(X_t) \xrightarrow[N \rightarrow \infty]{} \int dx \pi(x) O(x) = \langle O \rangle_\pi \quad (113)$$

となる。これが、Metropolis–Hastings 法で望みの期待値を求められる理由である。

6.7 自己相関と有効サンプル数

Monte Carlo 法で期待値を推定するときは、互いに独立な標本を用いるのが理想である。独立な標本 $x_i \sim p(x)$ が N 個あれば、標本平均 $\bar{O} = N^{-1} \sum_{i=1}^N O(x_i)$ の分散は

$$\text{Var}(\bar{O}) = \frac{1}{N} \text{Var}(O) \quad (114)$$

と評価できる。統計的な標本調査と同様に、標本間に相関があるにもかかわらず独立であるとみなすと、統計誤差を過小評価する可能性がある。

MCMC が生成するのは、定常分布が $p(x)$ であっても、一般には互いに相関した標本の系列である。ただし、標本間の相関を共分散として取り込めば、標本平均の統計誤差を評価できる。さらに、独立標本を用いる場合と比べて、MCMC がどれだけ非効率になっているかも定量化できる。系列 $O_i = O(x_i)$ に対し

$$\Gamma(t) = \langle (O_i - \langle O \rangle)(O_{i+t} - \langle O \rangle) \rangle, \quad \rho(t) = \frac{\Gamma(t)}{\Gamma(0)} \quad (115)$$

とする。積分自己相関時間を

$$\tau_{\text{int}} = \frac{1}{2} + \sum_{t=1}^{\infty} \rho(t) \quad (116)$$

と定義する。独立標本に対する統計的非効率性を

$$g = 2\tau_{\text{int}} = 1 + 2 \sum_{t=1}^{\infty} \rho(t) \quad (117)$$

と書けば、平均の分散は概ね

$$\text{Var}(\bar{O}) \simeq \frac{g}{N} \text{Var}(O) = \frac{2\tau_{\text{int}}}{N} \text{Var}(O) \quad (118)$$

である。従って有効独立サンプル数は

$$N_{\text{eff}} \simeq \frac{N}{g} = \frac{N}{2\tau_{\text{int}}}. \quad (119)$$

サンプラーの性能は 1 ステップの速さではなく、単位計算時間あたりの N_{eff} で比較すべきである。

上の分散公式は、標本平均の分散を直接展開すると得られる。

$$\text{Var}(\bar{O}) = \frac{1}{N^2} \sum_{i,j=1}^N \Gamma(|i-j|) \quad (120)$$

$$= \frac{\Gamma(0)}{N} \left[1 + 2 \sum_{t=1}^{N-1} \left(1 - \frac{t}{N} \right) \rho(t) \right]. \quad (121)$$

N が自己相関の長さより十分大きければ $(1 - t/N) \simeq 1$ として

$$\text{Var}(\bar{O}) \simeq \frac{2\tau_{\text{int}}}{N} \Gamma(0) \quad (122)$$

となる。相関した N 配位は、独立な N 配位と同じ情報量を持たない。また τ_{int} は観測量ごとに異なるため、ブラケットで短い自己相関が得られても、トポロジカルチャージや長距離相関関数でも短いとは限らない。

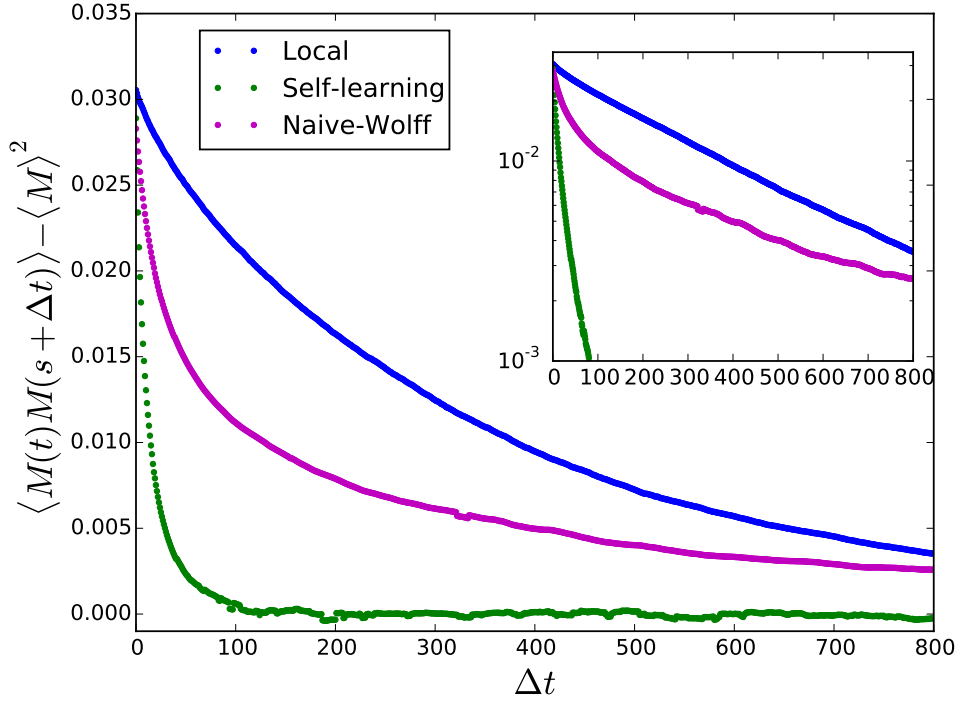


図 6: Self-learning Monte Carlo (SLMC) における自己相関関数の減衰 [18, Fig. 3]。2 次元正方格子 ($L = 40$) の臨界温度で、強磁性秩序変数 $M = N^{-1}|\sum_i S_i|$ の自己相関を比較している。青は局所更新、緑は self-learning 更新、紫は naive Wolff 型更新であり、右上は同じデータの片対数表示である。この条件では self-learning 更新の自己相関時間は局所更新より約 24 倍短い。

7 ハイブリッド・モンテカルロ法

HMC 法は、第 6 節で説明した Metropolis–Hastings 法の枠組みを保ちながら、提案の作り方を仮想的なハミルトン力学へ置き換えたものと理解できる。局所的な乱数提案の代わりに、作用の勾配を使って配位空間を長い距離だけ進む。Duane らは 1987 年、この方法を格子場理論へ導入した [17]。Metropolis 法、Metropolis–Hastings 法、HMC 法は、提案分布を改善しながら Metropolis–Hastings 補正で目標分布を守ると同じ系列にある。

7.1 仮想的なハミルトン力学

まず、 N 個の格子自由度 ϕ_n に対する経路積分

$$Z = \int \mathcal{D}\phi e^{-S[\phi]}, \quad \langle O \rangle = \frac{1}{Z} \int \mathcal{D}\phi O[\phi] e^{-S[\phi]} \quad (123)$$

から出発する。各 ϕ_n に対して仮想的な運動量 π_n を 1 つずつ導入し、規格化されたガウス積分

$$1 = \prod_{n=1}^N \left(\int_{-\infty}^{\infty} \frac{d\pi_n}{\sqrt{2\pi}} e^{-\pi_n^2/2} \right) \quad (124)$$

を経路積分へ挿入する。この操作は 1 を掛けただけなので、元の場の理論を変えない。 $p = (\pi_1, \dots, \pi_N)^T$ と書き、より一般に正定値行列 M を用いてガウス分布の共分散を調整すると、拡張された分配関数は

$$\begin{aligned} Z &= \frac{1}{(2\pi)^{N/2} (\det M)^{1/2}} \int \mathcal{D}\phi d^N p \exp \left[-S[\phi] - \frac{1}{2} p^T M^{-1} p \right] \\ &= \frac{1}{(2\pi)^{N/2} (\det M)^{1/2}} \int \mathcal{D}\phi d^N p e^{-H_{\text{HMC}}(\phi, p)} \end{aligned} \quad (125)$$

となる。ここで

$$H_{\text{HMC}}(\phi, p) = S(\phi) + \frac{1}{2} p^T M^{-1} p \quad (126)$$

を HMC ハミルトニアンと呼ぶ。 S が位置エネルギー、 $p^T M^{-1} p / 2$ が運動エネルギーに相当する。ただし、 H_{HMC} は元の物理系の実時間発展を生成するハミルトニアンではない。目標分布から効率よく配位を生成するために計算上導入した仮想的なハミルトニアンである。従って p と、これから導入する仮想時間 τ も計算上の変数であり、物理的な運動量や実時間ではない。

拡張位相空間の確率密度は

$$\rho(\phi, p) = \frac{1}{Z_{\text{ext}}} e^{-H_{\text{HMC}}(\phi, p)} \quad (127)$$

である。 p を積分すれば ϕ に依存しない定数になるため、場の周辺分布は元の $e^{-S[\phi]}$ のままである。従って、場だけに依存する観測量について

$$\langle O \rangle = \frac{\int \mathcal{D}\phi d^N p O[\phi] e^{-H_{\text{HMC}}(\phi, p)}}{\int \mathcal{D}\phi d^N p e^{-H_{\text{HMC}}(\phi, p)}} \quad (128)$$

が成り立つ。経路積分の期待値を、仮想的な古典力学系の正準分布における期待値として書き直したことになる。

この HMC ハミルトニアンから、仮想時間 τ に関するハミルトン方程式

$$\frac{d\phi}{d\tau} = M^{-1} p, \quad \frac{dp}{d\tau} = -\frac{\partial S}{\partial \phi} \quad (129)$$

を得る。厳密な時間発展では

$$\frac{dH_{\text{HMC}}}{d\tau} = \frac{\partial S^T}{\partial \phi} \frac{d\phi}{d\tau} + (M^{-1} p)^T \frac{dp}{d\tau} = 0 \quad (130)$$

となり、 H_{HMC} は保存される。また Hamilton 流は Liouville の定理により位相空間体積を保存し、時間反転に対して可逆である。従って 1 本の厳密な軌道は、 $H_{\text{HMC}}(\phi, p) = E$ という等エネルギー面上を動くミクロカノニカルな運動である。

このことと期待値の関係を明確にしておこう。状態密度とエネルギー E 上の平均を

$$\Omega(E) = \int \mathcal{D}\phi d^N p \delta(H_{\text{HMC}}(\phi, p) - E), \quad (131)$$

$$\langle O \rangle_E = \frac{1}{\Omega(E)} \int \mathcal{D}\phi d^N p O[\phi] \delta(H_{\text{HMC}}(\phi, p) - E) \quad (132)$$

と定義する。運動がそのエネルギー面上でエルゴード的なら、ミクロカノニカル平均は長時間平均

$$\langle O \rangle_E = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T d\tau O[\phi(\tau)] \quad (133)$$

から求められる。一方、元の経路積分が与える正準期待値は

$$\langle O \rangle = \frac{\int dE e^{-E} \Omega(E) \langle O \rangle_E}{\int dE e^{-E} \Omega(E)} \quad (134)$$

であり、ミクロカノニカル平均をエネルギーごとに正準重み $e^{-E} \Omega(E)$ で足し合わせたものになっている。従って、ハミルトン運動を利用して期待値を求められるが、1 本の固定エネルギー軌道だけで

元の正準期待値が得られるわけではない。HMC 法では各更新の冒頭で p をガウス分布から再生成してエネルギー面を移り、正しい重みで異なる軌道を選ぶ。厳密なハミルトン方程式を解ければ、各軌道上で H_{HMC} と $e^{-H_{\text{HMC}}}$ は変化せず、体積も保存されるため、軌道の始点と終点を入れ替えても確率密度は変わらない。

ここで ϕ は 1 つの格子点の場ではなく、全格子点・全内部自由度を並べた高次元ベクトルであり、 p も同じ次元を持つ。HMC 法の仮想時間発展は $\nabla S(\phi)$ を通して ϕ の全成分を同時に動かし、受理されたときには配位全体を $\phi \rightarrow \phi'$ と一括して置き換える。したがって、作用や力が局所的に定義されていても、HMC 法の 1 回の提案は大域的更新である。実際の数値積分には、可逆かつ位相空間体積を保存するリープフロッグ積分を用い、有限刻み幅によるエネルギー誤差を軌道末端の Metropolis 補正で取り除く。具体的な更新式と受容率は付録 A33 に示す。

7.2 HMC 法の 1 更新

前小節で導入した仮想的なハミルトン力学を、1 回のマルコフ更新としてまとめる。

Algorithm 1 HMC 法の 1 更新

- 1: 現在の場 ϕ を保存する
 - 2: $p \sim \mathcal{N}(0, M)$ を生成する
 - 3: $H_{\text{HMC}}^{\text{old}} = S(\phi) + \frac{1}{2}p^T M^{-1}p$ を計算する
 - 4: 全場成分にリープフロッグ積分を L 回適用し (ϕ', p') を得る
 - 5: $H_{\text{HMC}}^{\text{new}} = S(\phi') + \frac{1}{2}p'^T M^{-1}p'$ を計算する
 - 6: 確率 $\min(1, e^{-\Delta H_{\text{HMC}}})$ で配位全体 ϕ' を採用する
 - 7: 棄却なら保存した ϕ を再び記録する
-

最後の、棄却時に同じ配位をもう一度記録する手順には注意が要る。棄却配位をデータから捨てることマルコフ連鎖を変え、偏りを導入する。

7.3 自由スカラー場の HMC 力

HMC 法で配位を動かすには作用の勾配が必要である。まず自由スカラー場で具体形を求める。式 (31) の $\lambda_0 = 0$ に対し

$$\frac{\partial S}{\partial \varphi_n} = \sum_{\mu} (2\varphi_n - \varphi_{n+\hat{\mu}} - \varphi_{n-\hat{\mu}}) + (am_0)^2 \varphi_n. \quad (135)$$

力は $-\partial S / \partial \varphi_n$ である。これは格子ラプラシアンと質量項であり、局所的に計算できる。

HMC では、格子場 φ_n を古典力学の座標、導入した仮想運動量 p_n をその共役運動量とみなす。このとき Euclid 作用 $S[\varphi]$ は、HMC で導入した古典系のポテンシャルエネルギーの役割を果たす。単位質量行列 $M = 1$ を選ぶと、HMC のハミルトン方程式は

$$\frac{d\varphi_n}{d\tau} = p_n, \quad (136)$$

$$\frac{dp_n}{d\tau} = \sum_{\mu} (\varphi_{n+\hat{\mu}} + \varphi_{n-\hat{\mu}} - 2\varphi_n) - (am_0)^2 \varphi_n = (\Delta_{\text{lat}} \varphi)_n - (am_0)^2 \varphi_n \quad (137)$$

となる。ここで τ は HMC の仮想時間であり、

$$(\Delta_{\text{lat}} \varphi)_n = \sum_{\mu} (\varphi_{n+\hat{\mu}} + \varphi_{n-\hat{\mu}} - 2\varphi_n) \quad (138)$$

と定義した。2 式を組み合わせると

$$\frac{d^2 \varphi_n}{d\tau^2} = (\Delta_{\text{lat}} \varphi)_n - (am_0)^2 \varphi_n \quad (139)$$

を得る。これは、古典的な自由スカラー場のハミルトニアン

$$H_{\text{cl}} = \int d^d x \left[\frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \right] \quad (140)$$

から得られる運動方程式

$$\frac{\partial^2 \phi}{\partial t^2} = \nabla^2 \phi - m^2 \phi \quad (141)$$

の格子版と同じ形をしている。すなわち、 $-\partial S / \partial \varphi_n$ は近接格子点との結合と質量項から生じる復元力であり、HMC 軌道はこの力を受ける古典的な結合振動子系として理解できる。

実際、 $\varphi_n(\tau) \propto e^{iq \cdot n} \varphi(q, \tau)$ と Fourier 変換すると、

$$\frac{d^2 \varphi(q, \tau)}{d\tau^2} = -\omega_q^2 \varphi(q, \tau), \quad (142)$$

$$\omega_q^2 = 4 \sum_{\mu} \sin^2 \left(\frac{q_{\mu}}{2} \right) + (am_0)^2 \quad (143)$$

となり、各運動量モードは古典的な調和振動子として仮想時間内を運動する。ただし、この対応は運動方程式の数学的な形についてのものである。格子点 n は Euclid 時空全体を走り、 p_n と τ はサンプリングのために導入した仮想変数である。従って HMC 軌道は元の量子場の実時間発展を表さない。

8 新しいサンプリングの必要性

8.1 HMC 法の計算コストと臨界減速

HMC 法は厳密な更新法だが、格子を細かくし、体積を大きくすると計算効率が悪化する。主な要因は次の通りである。

1. 格子体積が増えると力計算が高価になる。
2. 軽いクォークでは Dirac 演算子の条件数が悪化し、線形ソルバーが遅くなる。
3. $a \rightarrow 0$ で長距離モードが遅くなり、自己相関が増える。
4. トポロジカルセクター間の障壁が高くなり、トポロジカル・フリージングが起きる。

1 軌道の計算量は模式的に

$$C_{\text{traj}} \sim N_{\text{step}} (C_{\text{gauge}} + N_{\text{solve}} C_{D-1}) \quad (144)$$

と書ける。 N_{step} は分子動力学の刻み数、 C_{D-1} は Dirac 演算子を含む線形方程式を 1 回解く費用である。動的フェルミオン計算では C_{D-1} が支配的になりやすい。刻み幅を小さくすれば受容率は上がるが、 N_{step} も増えるため、1 軌道の受容率だけを最大にしても全体は速くならない。

連続極限に近づくと、観測量 O の自己相関時間はしばしば

$$\tau_{\text{int}}(O) \propto \xi^{z_O} \quad (145)$$

のように相関長 ξ とともに増大する。 z_O はアルゴリズムと観測量に依存する動的臨界指数である。自己相関関数を単純化して

$$\rho(\tau) \simeq \exp \left(-\frac{\tau}{\tau_{\text{exp}}} \right) \quad (146)$$

と書けば、臨界点へ近づくほど減衰時間 τ_{exp} が長くなり、曲線は横に広がる。

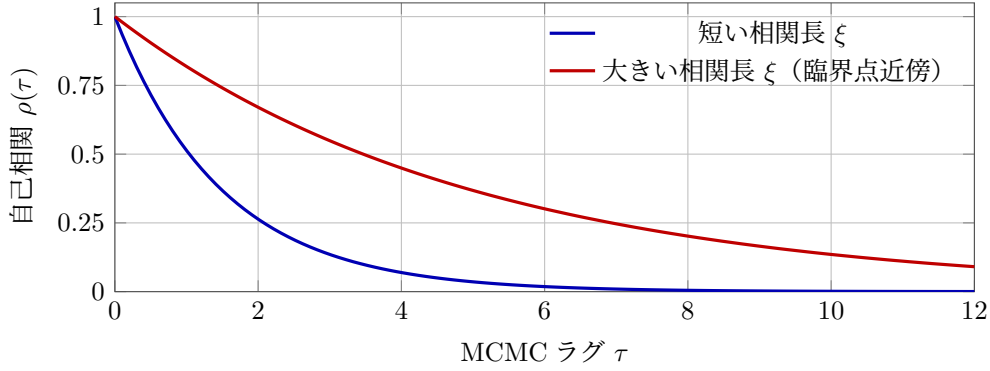


図 7: 臨界減速の模式図。相関長が大きくなると自己相関関数 $\rho(\tau)$ の減衰が遅くなり、 τ_{int} も増大する。同じ数の配位を保存しても、独立な情報量は少なくなる。横軸の τ は HMC 軌道内の仮想時間ではなく、保存配位間の MCMC ラグである。

従って独立な情報を 1 つ得る費用は、およそ $C_{\text{traj}}\tau_{\text{int}}(O)$ で決まる。機械学習を導入する動機は、1 回の力計算の高速化にとどまらず、典型集合の大域構造を学んで相関時間そのものを短くすることにある。

8.2 全結合ニューラルネット

まず、教師あり学習で使うデータを $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$ と書く。各 $(\mathbf{x}_i, \mathbf{y}_i)$ は、入力とそれに対応する出力の組である。例えば物理実験で、時刻 x_i とその時刻に測った移動距離 y_i を組にした表を考えればよい。このデータから、未知の時刻 x に対する距離 y を予測する関数を作りたい。

ニューラルネットは、そのために用意するパラメータ付きの変分近似関数 $f_\theta: \mathbb{R}^{d_0} \rightarrow \mathbb{R}^{d_L}$ である。 θ は重みやバイアスなどの調整可能なパラメータをまとめた記号であり、最初から正しい関数 f を知っているわけではない。関数族 $\{f_\theta\}$ の中から、各データ点で $f_\theta(\mathbf{x}_i) \simeq \mathbf{y}_i$ となる θ を選ぶ。具体的には

$$\hat{\mathcal{L}}(\theta) = \frac{1}{N} \sum_{i=1}^N \ell(f_\theta(\mathbf{x}_i), \mathbf{y}_i), \quad \theta_* = \arg \min_{\theta} \hat{\mathcal{L}}(\theta) \quad (147)$$

のように、予測とデータのずれを測る損失関数を最小化する。この意味でニューラルネットは、データ点を補間し、未知の入力にも使える関数をフィッティングする道具である。

これは、試行関数族を定め、そのパラメータを最適化するという意味で、広い意味での変分法である。ただし、方程式を解析的に解いて変分パラメータを決めるのではなく、有限個の観測例や標本から作った損失関数によって決めるため、**データ駆動型の変分法**と見なせる。生成モデルやサンプリングへの応用では明示的な正解ラベルがない場合もあるが、対数尤度、KL divergence、作用などから作る損失を標本上で評価し、同じように θ を調整する。

最も基本的なニューラルネットとして、全結合フィードフォワードネットワークを考える。入力を縦ベクトル $\mathbf{x}^{(0)} = \mathbf{x} \in \mathbb{R}^{d_0}$ とし、各層で

$$\begin{aligned} \mathbf{x}^{(\ell)} &= \sigma\left(W^{(\ell)}\mathbf{x}^{(\ell-1)} + \mathbf{b}^{(\ell)}\right), \quad \ell = 1, \dots, L-1, \\ f_\theta(\mathbf{x}) &= W^{(L)}\mathbf{x}^{(L-1)} + \mathbf{b}^{(L)} \end{aligned} \quad (148)$$

と定義する。 $W^{(\ell)}$ は重み行列、 $\mathbf{b}^{(\ell)}$ はバイアスであり、パラメータ θ はこれらすべてを表す。「全結合」とは、ある層の各成分が次の層のすべての成分へ重みを介して接続されることを意味する。活性化関数 σ を入れずに線形変換だけを重ねても、全体は 1 つのアフィン変換にまとめられる。従って複雑な非線形関数を表す能力は、層数だけでなく σ の非線形性から生じる。

1 つの隠れ層を持つスカラー出力のネットワークは

$$f_\theta(\mathbf{x}) = \sum_{r=1}^m a_r \sigma(\mathbf{w}_r^T \mathbf{x} + b_r) + c \quad (149)$$

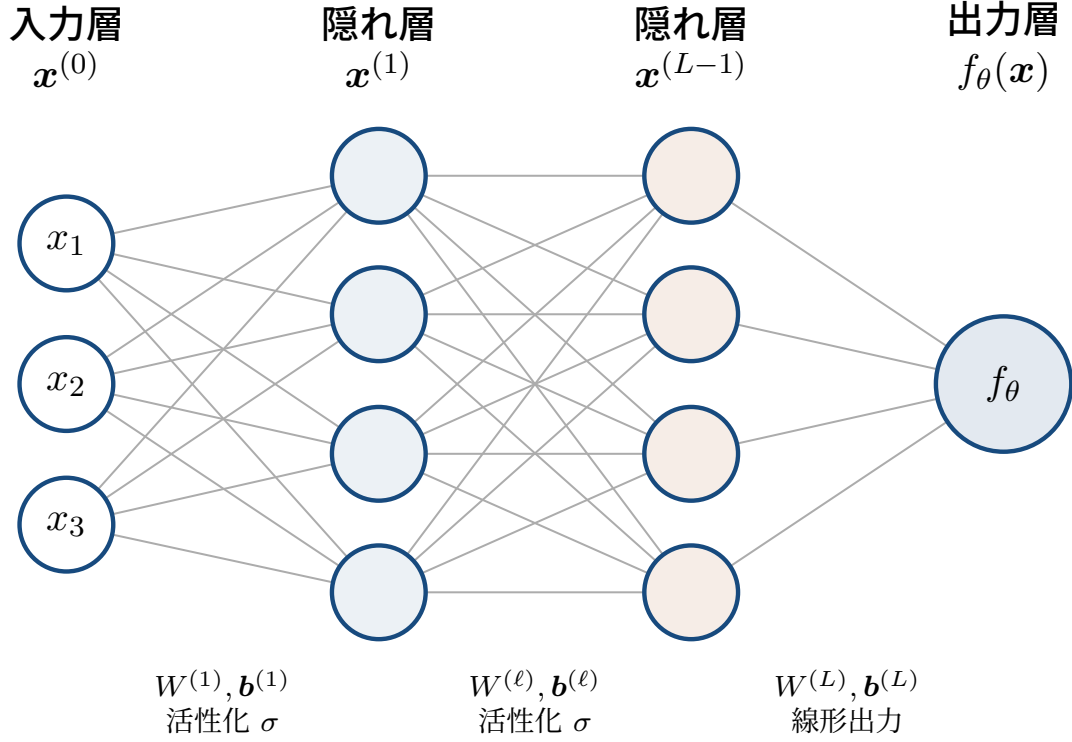


図 8: 全結合フィードフォワードネットワークの模式図。隣接する層の各ノードはすべて結ばれ、各線に対応する重みと各層のバイアスが学習される。隠れ層ではアフィン変換の後に活性化関数 σ を適用する。

と書ける。万能近似定理 (universal approximation theorem、普遍近似定理ともいう) によれば、シグモイド型活性化関数を用いると、コンパクト領域 $K \subset \mathbb{R}^{d_0}$ 上の任意の連続関数 $f: K \rightarrow \mathbb{R}$ と任意の $\varepsilon > 0$ に対し、隠れ層の幅 m とパラメータを適切に選んで

$$\sup_{\mathbf{x} \in K} |f(\mathbf{x}) - f_{\theta}(\mathbf{x})| < \varepsilon \quad (150)$$

とできる [19]。すなわち、十分に幅の広い全結合ネットワークは、連続関数を任意の精度で近似する表現能力を持つ。

ただし、この定理は近似するパラメータの存在を述べるだけであり、必要な幅、学習データ数、最適化によってそのパラメータへ到達できるか、未知の入力へ一般化できるかを保証しない。格子場の全自由度をそのまま全結合するとパラメータ数も急増し、局所性や対称性も自動では保たれない。この講義で全結合ネットワークを導入するのは、後述する正規化フローの内部で、可逆変換を決めるスケールや平行移動を柔軟に表現する部分写像として利用するためであって、万能近似定理だけでサンプリング問題が解けるからではない。

8.3 学習モデルと厳密性の役割分担

学習モデルの役割は、物理量を直接予測することではなく、現在の配位 x から次の候補 x' を作る proposal を改善することである。学習した提案核 $q_{\theta}(x'|x)$ で候補を作り、

$$A_{\theta}(x \rightarrow x') = \min \left[1, \frac{\pi(x')q_{\theta}(x|x')}{\pi(x)q_{\theta}(x'|x)} \right]$$

で受容・棄却する。受容率は正確な作用と提案確率から計算する。

提案確率を正しく評価でき、分布の台などの条件を満たせば、学習モデルが不完全でも Metropolis–Hastings 補正後の定常分布は π に保たれる。近似誤差は主として受容率や自己相関を悪化させるが、

そのまま物理量のバイアスにはならない。すなわち、学習モデルが効率のよい候補を作り、確率論的な受容規則が定常分布を保証する。

ただし、有限標本誤差、熱化不足、エルゴード性の問題は別に検証する必要がある。独立提案の場合の式と成立条件、誤差の分類は付録 A32 にまとめる。

9 正規化フローと Metropolis 補正

9.1 変数変換に基づく生成モデル

生成モデルを導入する前に、経路積分そのものの変数変換を考える。格子上の場をまとめて ϕ と書くと

$$Z = \int \mathcal{D}\phi e^{-S[\phi]} \quad (151)$$

である。可逆な変数変換 $\phi = f(z)$ と、その関数ヤコビアン

$$J_f(z) = \frac{\partial f(z)}{\partial z} \quad (152)$$

を導入すれば

$$\begin{aligned} Z &= \int \mathcal{D}z |\det J_f(z)| e^{-S[f(z)]} \\ &= \int \mathcal{D}z e^{-S_f[z]}, \quad S_f[z] = S[f(z)] - \log |\det J_f(z)| \end{aligned} \quad (153)$$

となる。従って変数変換は、作用の引き戻し $S[f(z)]$ だけでなく、測度の変化 $-\log |\det J_f|$ も新しい作用へ加える。

基底作用 $S_0[z]$ は、対応する分布 $r(z) \propto e^{-S_0[z]}$ から標本を容易に生成できるように選ぶ。例えば運動項がなく、 $S_0[z] = \sum_n s_0(z_n)$ と各格子点の局所項の和に分解できれば、各 z_n を独立にサンプルできる。もし

$$S[f(z)] - \log |\det J_f(z)| = S_0[z] + \text{const.} \quad (154)$$

を満たす写像 f が分かれば、 $r(z) \propto e^{-S_0[z]}$ から独立な z を生成し、 $\phi = f(z)$ と変換するだけで相互作用する理論の厳密標本が得られる。このように相互作用する測度を簡単な測度へ移す変換を、広い意味で trivializing map と呼ぶ。

ヤコビアンは単なる規格化因子ではない。 J_f が場に依存すれば $-\log |\det J_f|$ も場に依存する作用項となり、変換された基底作用とともに、隣接点の結合や運動項を含む非自明な構造を作ったり打ち消したりする。運動項が常にヤコビアン単独から生じるわけではなく、 $S_0[f^{-1}(\phi)]$ とヤコビアン項の組合せで変換後の作用全体が決まることに注意する。

Nicolai 写像 この発想の重要な先例が、超対称理論における Nicolai 写像である [20]。フェルミオンを積分した後の理論を模式的に

$$Z_{\text{SUSY}} = \int \mathcal{D}\phi \det M[\phi] e^{-S_B[\phi]} \quad (155)$$

と書く。ボソン場の非線形変換 $\eta = \mathcal{N}[\phi]$ が

$$S_B[\phi] = \frac{1}{2} \int dx \eta(x)^2, \quad \det \frac{\delta \mathcal{N}}{\delta \phi} = \det M[\phi] \quad (156)$$

を満たせば、関数ヤコビアンがフェルミオン行列式を再現し、経路積分は形式的に

$$Z_{\text{SUSY}} = \int \mathcal{D}\eta \exp \left[-\frac{1}{2} \int dx \eta(x)^2 \right] \quad (157)$$

というガウス測度へ移る。例えば $\eta = \partial\phi + W'(\phi)$ のように微分を含む写像では、 $\eta^2/2$ を展開した項がボソンの運動項 $(\partial\phi)^2/2$ とポテンシャルを生み、写像のヤコビアンがフェルミオン行列式を与える。実際には逆写像の多価性、行列式の符号、大域的な存在条件があり、上式は Nicolai 写像の役割を示す模式的な関係である。

Lüscher の trivializing map 超対称性を仮定しない格子ゲージ理論でも、相互作用する変数を互いに独立な変数へ写すという同じ目標を立てられる。Lüscher は、場の空間における flow 方程式を積分して trivializing map を系統的に構成し、Wilson flow が Wilson ゲージ作用に対する近似的な trivializing map を与えること、さらに変換後の変数で HMC 法を実行する方向を示した [21]。

ニューラルネットによる近似 厳密な trivializing map を解析的に求めることは一般には難しい。そこで可逆写像を f_θ でパラメータ化し、式 (154) を近似するようニューラルネットを学習する。正規化フロー (normalizing flow) は、簡単な基底分布

$$z \sim r(z), \quad r(z) = \mathcal{N}(0, I) \quad (158)$$

から可逆写像

$$\phi = f_\theta(z), \quad z = f_\theta^{-1}(\phi) \quad (159)$$

で場の配位を生成する。そのモデル密度は

$$q_\theta(\phi) = r(f_\theta^{-1}(\phi)) \left| \det \frac{\partial f_\theta^{-1}(\phi)}{\partial \phi} \right|, \quad (160)$$

$$\log q_\theta(f_\theta(z)) = \log r(z) - \log \left| \det \frac{\partial f_\theta(z)}{\partial z} \right| \quad (161)$$

で評価できる。近似写像が完全に trivialize しなくても、小節 8.3 で説明した MH 補正を加えれば最終的な定常分布を厳密に保てる。格子場理論におけるこの基本形は Albergo ら [33] によって示された。

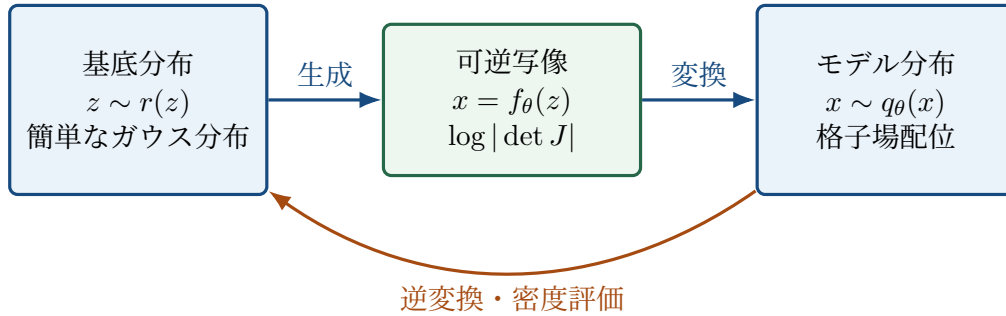


図 9: 正規化フローの 3 要素。生成だけでなく、逆変換とヤコビアンを通じて $q_\theta(x)$ を評価できることが MH 補正に必要である。

9.2 逆 KL 学習と Metropolis–Hastings 補正

正規化フローはスカラー場だけでなく、ゲージ場にも構成できる。ゲージ対称性を層ごとに保つ $U(1)$ および $SU(N)$ のフローが提案されており [37, 38]、4 次元の $N_f = 2$ QCD、すなわち動的クォークを含む理論でも、フローを用いた数値的実例が報告されている [40, 41]。ゲージ共変な写像の条件と具体的なネットワーク構成は、付録 A19 にまとめる。

目標分布の規格化定数 Z を知らなくても、逆 KL ダイバージェンス

$$D_{\text{KL}}(q_\theta \parallel \pi) = \int dx \, q_\theta(x) \log \frac{q_\theta(x)}{\pi(x)} \quad (162)$$

$$= \mathbb{E}_{x \sim q_\theta} [\log q_\theta(x) + S(x)] + \log Z \quad (163)$$

を最小化できる。 $\log Z$ は θ に依存しないので学習に不要である。基底分布からの標本 z_i を用いれば損失関数は

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_i [\log q_\theta(f_\theta(z_i)) + S(f_\theta(z_i))]. \quad (164)$$

これは既存の QCD 配位を教師データとして必要としない自己学習である。

勾配は基底分布からの標本を固定した再パラメータ化によって計算できる。

$$\nabla_\theta \mathcal{L} = \mathbb{E}_{z \sim r} \left[\nabla_\theta \log q_\theta(f_\theta(z)) + \frac{\partial S(x)}{\partial x} \bigg|_{x=f_\theta(z)} \frac{\partial f_\theta(z)}{\partial \theta} \right]. \quad (165)$$

第 2 項には作用の勾配が現れ、HMC 法の力と共通する計算要素を持つ。

逆 KL $D_{\text{KL}}(q||\pi)$ と、教師配位 $x \sim \pi$ を用いる順 KL $D_{\text{KL}}(\pi||q)$ は性質が異なる。

$$D_{\text{KL}}(\pi||q_\theta) = \text{const.} - \mathbb{E}_{x \sim \pi} [\log q_\theta(x)]. \quad (166)$$

順 KL は目標分布からの標本を必要とするが、確率のあるモードを取りこぼすと強く罰する。逆 KL は作用だけで学習できる一方、 q が訪れないモードを直接は罰しない。どちらを使うかは、教師配位の有無と分布全体を被覆する必要性に依存する。モード被覆を検証する際の注意点は付録でまとめる。

学習したフローを目標分布からのサンプリングに用いるには、その出力を独立提案として Metropolis–Hastings 法で補正する。すなわち、現在点に依存せず $x' \sim q_\theta(x')$ をフローから提案する。式 (112) で $q(x'|x) = q_\theta(x')$ だから

$$A(x \rightarrow x') = \min \left[1, \frac{\pi(x')q_\theta(x)}{\pi(x)q_\theta(x')} \right]. \quad (167)$$

作用で書けば

$$\log R = -S(x') + S(x) + \log q_\theta(x) - \log q_\theta(x'), \quad A = \min(1, e^{\log R}). \quad (168)$$

ここでも Z は消える。

フローが不完全でも、 $q_\theta(x) > 0$ である領域が目標分布の台を覆い、密度が正しく評価できれば、MH 補正後の定常分布は厳密に π になる。学習誤差は受容率と自己相関へ影響するが、物理量の系統誤差にはしない。

この主張は詳細釣り合いから直接確認できる。 $x \neq x'$ に対して

$$\pi(x)q_\theta(x')A(x \rightarrow x') \quad (169)$$

$$= \min [\pi(x)q_\theta(x'), \pi(x')q_\theta(x)] \quad (170)$$

$$= \pi(x')q_\theta(x)A(x' \rightarrow x). \quad (171)$$

これは式 (290) の独立提案版である。 $q_\theta = \pi$ なら受容比は常に 1 で、各ステップが目標分布からの独立標本になる。 q_θ がずれると棄却が生じ、棄却時には同じ状態が連続するため自己相関が生まれる。

重要度重み

$$w(x) = \frac{e^{-S(x)}}{q_\theta(x)} \quad (172)$$

を用いれば、受容比は $R = w(x')/w(x)$ である。 $\log w$ の幅が狭いほど受容率は高くなりやすいが、平均受容率だけでは稀なモードの欠落は検出できない。したがって重み分布、観測量、トポロジカルセクターの被覆を同時に調べる必要がある。

9.3 研究例：ゲージ同変フローとトポロジーの凍結

正規化フローが実際のサンプリングをどう変えるかを見るため、Kanwar らによる 2 次元 $U(1)$ 格子ゲージ理論の研究を取り上げる [37]。16 × 16 格子で $\beta = 1, \dots, 7$ を調べ、24 層のゲージ同変な coupling layer からなるフローを学習し、フローの出力を独立提案型 Metropolis 法で補正した。各 β についてフロー、HMC 法、Heat Bath 法でそれぞれ 10^5 配位の連鎖を生成している。

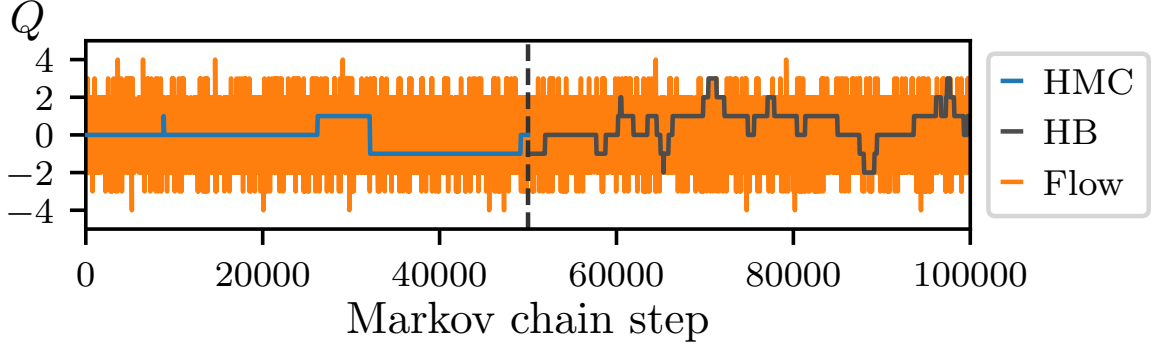


図 10: $\beta = 7$ におけるトポロジカル電荷 Q の履歴。HMC 法と Heat Bath 法では長時間同じトポロジカルセクターに留まるのに対し、ゲージ同変フローを用いた Metropolis 法は異なるセクターを頻繁に訪れる。Kanwar ら [37] (arXiv:2003.06413) の図 1。

履歴を眺めるだけでなく、トポロジカル電荷の積分自己相関時間 $\tau_{\text{int}}(Q)$ を比較すると差はさらに明確になる。 β を大きくして連続極限へ近づくにつれ、HMC 法と Heat Bath 法では $\tau_{\text{int}}(Q)$ が急増するが、フローでは増加が大幅に抑えられている。

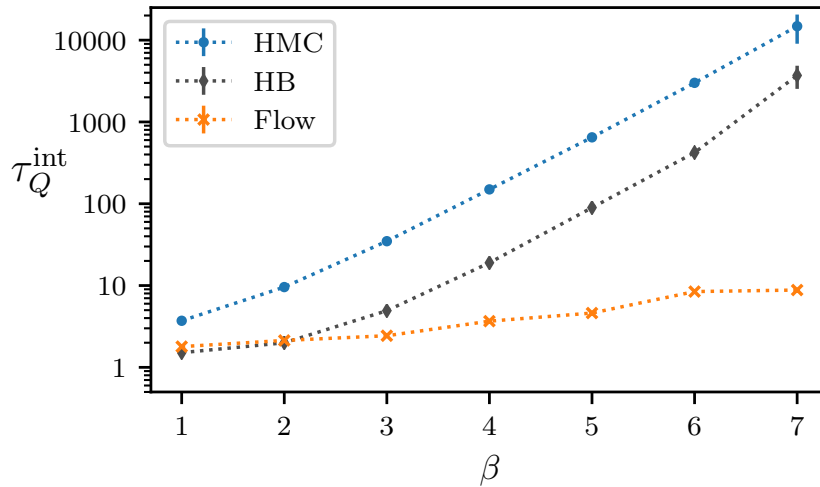


図 11: 16 × 16 格子における $\tau_{\text{int}}(Q)$ の β 依存性。 $\beta = 7$ ではフローが約 10、Heat Bath 法が約 4×10^3 、HMC 法が約 1.5×10^4 である。Kanwar ら [37] (arXiv:2003.06413) の図 4。

原論文の単一 CPU 比較では、フロー 1 提案と HMC 法 1 軌道の費用は約 10% 以内で同程度であり、 $\beta = 7$ のトポロジカル量では、フローを HMC 法の約 1500 倍、Heat Bath 法の約 200 倍効率的と見積もった。これは 2 次元 $U(1)$ 理論での proof of principle であり、4 次元 QCD の性能保証ではない。それでも、対称性を保つ学習提案と MH 補正が、正しさを保ったままトポロジーの凍結を緩和し得る実例である。

9.4 研究例：動的クォークを含む 4 次元 QCD

QCD では、ゲージ場 U だけでなく、動的クォークの行列式も目標分布へ含める必要がある。Abbott らは、 4^4 格子、 $N_f = 2$ 、 $\beta = 1$ 、 $\kappa = 0.1$ の 4 次元 QCD で、ゲージ場と擬フェルミオン場を同時に生

成するフローを実証した [40]。提案分布を

$$q(U, \phi) = q_m(U)q_c(\phi | U) \quad (173)$$

と分解し、marginal flow q_m で SU(3) ゲージ場を、conditional flow q_c でその背景ゲージ場に依存する擬フェルミオン場を生成する。後者がフェルミオン行列式の非局所な効果を学習する役割を担う。

1つのゲージ場につき 512 個の擬フェルミオン標本で行列式の重みを推定すると、学習後のモデルは $\text{ESS} \simeq 0.1$ を達成した。一方、Haar 一様分布から QCD へ直接再重み付けした場合は $\text{ESS} \lesssim 10^{-4}$ であった。フローから生成した 6.5×10^4 配位は、およそ 6.5×10^3 個の独立標本に相当し、同等統計の HMC 法と比べてプラケット、Polyakov loop、パイオン相関関数、トポロジカル電荷の分布が統計誤差内で整合した。

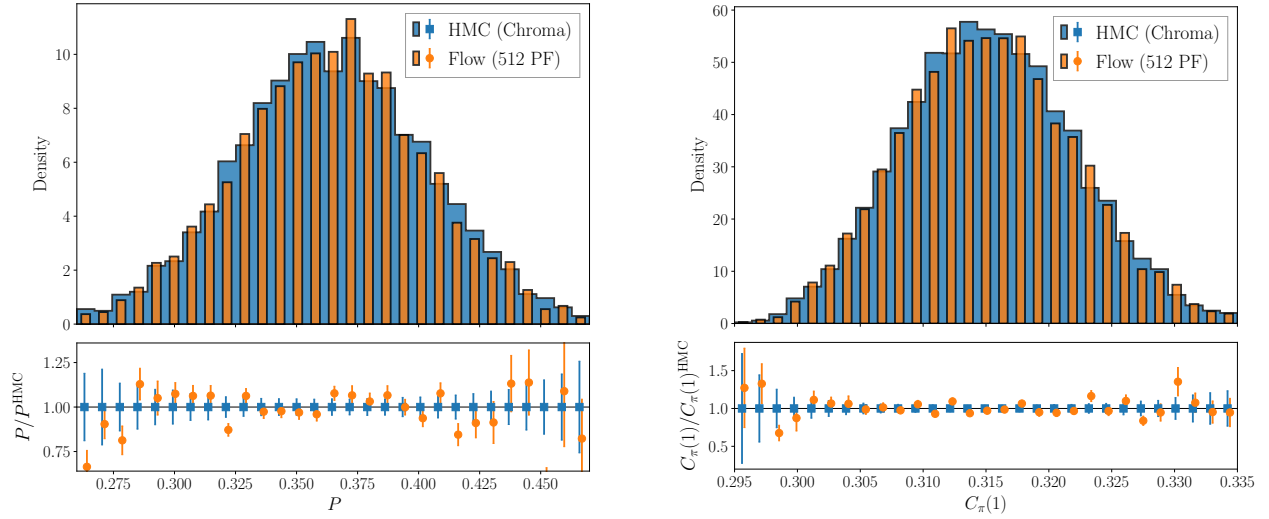


図 12: $N_f = 2$ の 4 次元 QCD における、HMC 法（青）とフロー（橙）の比較。左はプラケット、右は時刻差 1 のパイオン相関関数の分布であり、下段は HMC 法に対する比を示す。Abbott ら [40] (arXiv:2208.03832) の図 2 から抜粋。

これは現実的な格子間隔・体積で HMC 法を上回ったという結果ではない。非常に小さく粗い格子での proof of principle である。それでも、SU(3) ゲージ対称性だけでなく、動的クォークに由来する行列式まで含めて、正規化フローと再重み付けを QCD へ接続できることを示した点が重要である。

10 拡散模型

拡散模型は、複雑な分布から得た配位に少しずつ Gauss ノイズを加え、その逆向き過程をニューラルネットワークで学習する生成模型である [27, 28, 29]。図 13 の上段では、構造をもつ配位が段階的に乱され、最後には元の配位を目で識別できないノイズへ近づく。下段の生成過程はこの向きを逆にたどり、各時刻で残すべき構造を score から推定する。

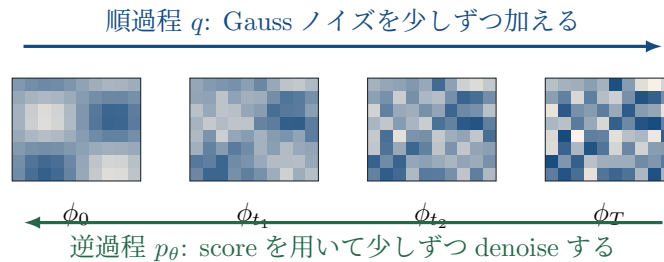


図 13: 拡散模型の順過程と逆過程。色は格子点上の場の値を表す。順過程では配位の構造を Gauss ノイズで徐々に壊し、逆過程ではノイズから新しい配位を生成する。Yang らの総説の図 2[29] を参考に、格子場の配位へ置き換えて再構成した。

目標分布 π から得た配位を ϕ_0 とし、 $t \in [0, T]$ で

$$\phi_t = \alpha_t \phi_0 + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \quad (174)$$

と加音する。 α_t を減少、 σ_t を増加させれば、 $t = T$ で $p_T(\phi_T)$ は扱いやすい Gauss 分布へ近づく。

時刻 t の分布の score $\nabla_{\phi_t} \log p_t(\phi_t)$ を $s_\theta(\phi_t, t)$ で近似する。式 (174) では条件付き score が $-\epsilon/\sigma_t$ なので、denoising score matching の損失は

$$\mathcal{L}_{\text{score}}(\theta) = \mathbb{E}_{t, \phi_0, \epsilon} \left[w(t) \left\| s_\theta(\phi_t, t) + \frac{\epsilon}{\sigma_t} \right\|^2 \right] \quad (175)$$

と書ける。学習後は Gauss ノイズから出発し、score を用いた逆向き SDE を T から 0 へ数値積分して配位を生成する。

正規化フローと異なり、写像全体を 1 つの可逆変換にする必要も、各層の Jacobian を計算する必要もない。その代わり生成には多数のネットワーク評価が必要であり、離散化誤差と score の学習誤差も入る。格子場へ用いる場合は並進対称性やゲージ対称性をネットワークに組み込み、生成配位をそのまま厳密標本とみなさないことが重要である。厳密な期待値が必要ななら、評価可能な提案確率を備えた補正、再重み付け、または別の厳密な MCMC 法との組み合わせが必要になる。

10.1 研究例：Lingxiao Wang らの拡散サンプラー

Wang, Aarts, Zhou は、拡散模型の逆向き確率過程を確率過程量子化の Langevin dynamics と対応づけ、2 次元 ϕ^4 理論で拡散模型を大域的提案として用いると、特に臨界点近傍で自己相関を短縮できることを示した [30]。この見方では、score $\nabla_{\phi} \log p_t(\phi)$ は、各 noise level で配位を確率の高い方向へ戻す Langevin drift である。

さらに Zhu, Aarts, Wang らは、物理パラメータで score を条件付け、Metropolis 補正付き annealed Langevin 法と組み合わせて、2 次元 compact $U(1)$ 格子ゲージ理論へ拡張した [31]。 $\beta = 1$ の教師配位で学習した模型を $\beta = 7$ へ移すと、HMC 法がほぼ $Q = 0$ に凍結する条件でも、拡散提案は複数のトポロジカルセクターを訪れた。

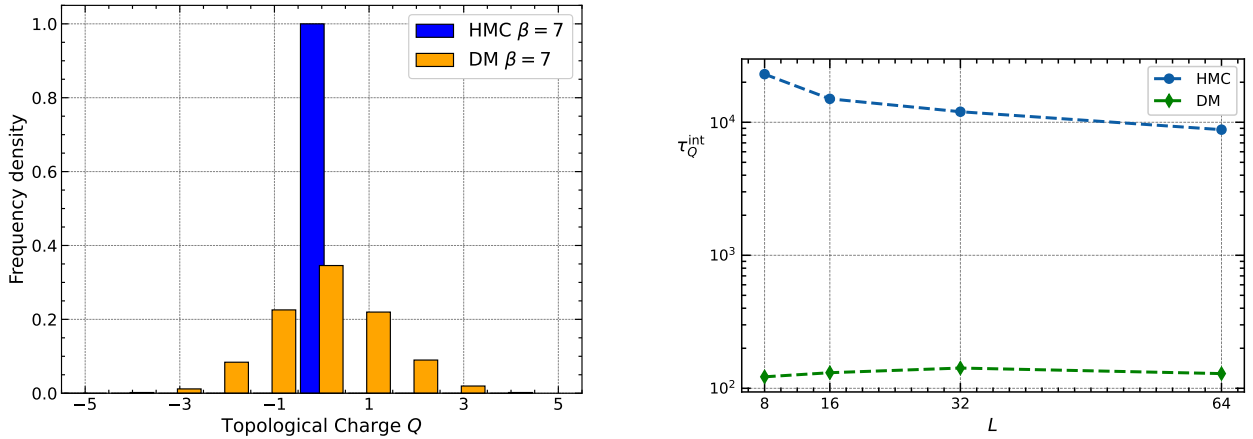


図 14: $\beta = 7$ における Lingxiao Wang らの physics-conditioned diffusion sampler。左は $L = 16$ のトポロジカル電荷分布で、HMC 法が $Q = 0$ 付近に凍結する一方、拡散模型は複数セクターを訪れる。右は $L = 8, 16, 32, 64$ の $\tau_{\text{int}}(Q)$ で、HMC 法の約 $8 \times 10^3 - 2 \times 10^4$ に対し、拡散模型は約 $1.2 - 1.4 \times 10^2$ である。Zhu ら [31] (arXiv:2502.05504) の図 3 右および図 4 左。

同じ計算で拡散提案の受容率は $L = 8$ の約 0.76 から $L = 64$ の約 0.45 まで緩やかに低下した。ただし、これは $\beta = 1$ で従来法が生成した教師 ensemble を必要とする。Wilson loop は $\beta = 11$ まで良好に転用できた一方、トポロジカル感受率の外挿は $\beta \gtrsim 5$ で破綻し始めたと報告されている。従って、セクター間移動の改善と、目標分布を有限統計で正確に再現することは分けて検証しなければならない。

10.2 研究例：自己学習拡散

通常の score model は、目標結合で別の Monte Carlo 法が作った教師配位を必要とする。富谷の SLDiffusion は、外部教師配位なしで拡散サンプラーを構成する方法である [32]。2 次元 compact XY 模型では、厳密標本を生成できる $\beta = 0$ から出発し、 $0 = \beta_0 < \beta_1 < \dots < \beta_K$ という梯子を上る。 β_k で score を学習して固定し、次の β_{k+1} を目標とする各 noise level の提案を MH 補正する。補正済み replay だけを次段階で使うため、学習データを $\beta = 0$ まで追跡できる。

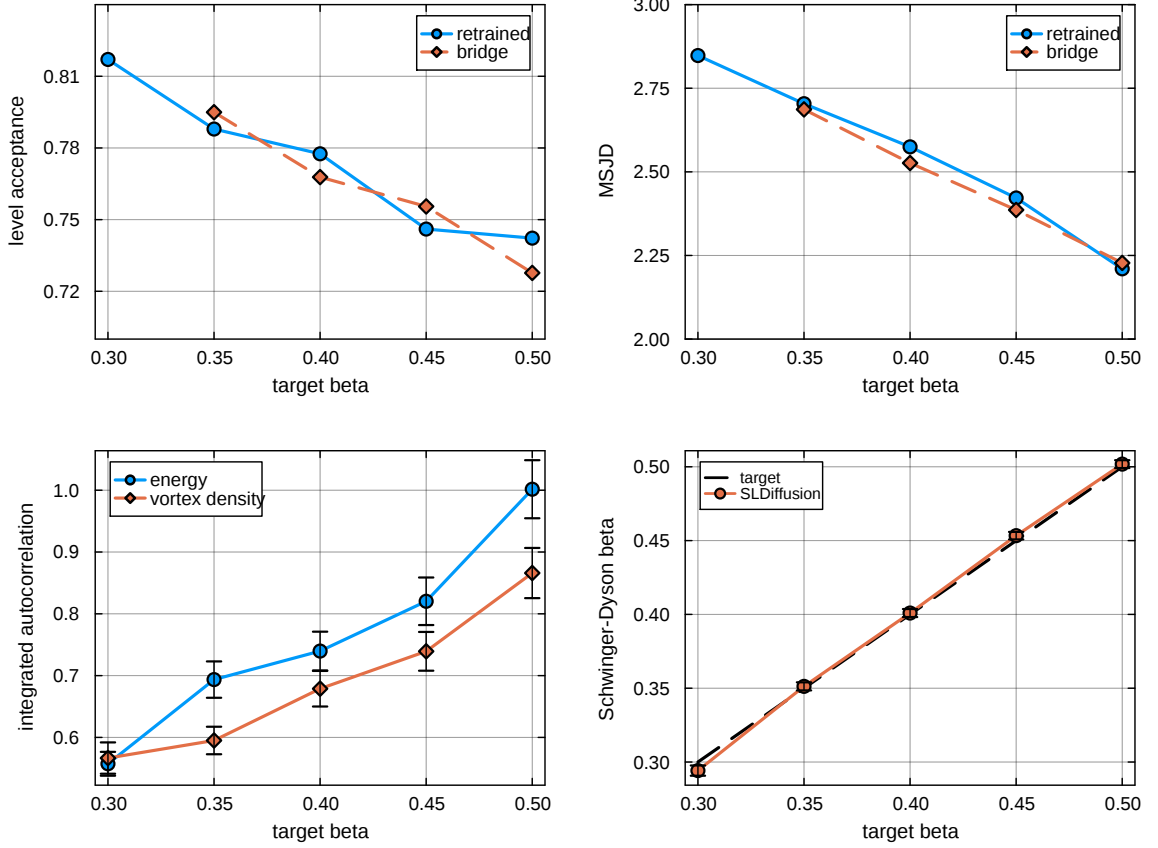


図 15: $L = 4$ の self-bootstrap β ladder。上段は noise level ごとの受容率と平均二乗移動距離 (MSJD)、下段左はエネルギー密度と渦密度の積分自己相関時間、下段右は Schwinger–Dyson 関係から推定した有効結合を示す。Tomiya[32] (arXiv:2607.12587) の図 2。

$L = 4$ では $\beta = 0.30$ から 0.50 まで外部教師配位なしで学習し、受容率は 0.817 から 0.742、エネルギー密度と渦密度の τ_{int} は 1.1 未満であった。Schwinger–Dyson 有効結合も目標 β と整合する。いずれもニューラルネットの出力そのものではなく、MH 補正済み連鎖上の結果である。

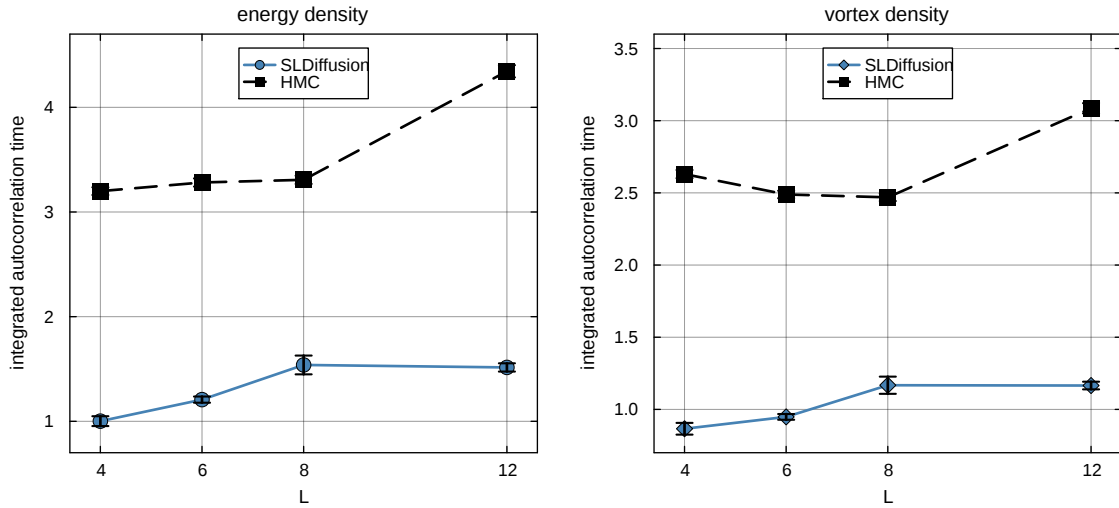


図 16: $\beta = 0.5$ における自己学習拡散と HMC 法の積分自己相関時間。各体積で replay を用いて再学習すると、 $L = 4, 6, 8, 12$ でエネルギー密度と渦密度の τ_{int} は 2 未満である。横軸の比較単位は保存 trajectory である。Tomiya[32] (arXiv:2607.12587) の図 5。

$\beta = 0.5$ では $L = 4, 6, 8, 12$ の 2 観測量が独立 HMC 計算と 1.35σ 以内で一致し、 $\tau_{\text{int}} < 2$ であった。ただし、自己学習拡散 1 trajectory は 32 回の全格子更新、HMC trajectory は長さ 1 であり、図は wall-clock 時間の優越を示さない。成果は、外部教師 ensemble なしに、MH 補正の厳密性を保って体積ごとの再学習を実証した点にある。

11 第 2 講のまとめ

第 2 講の要点

1. MCMC は規格化定数を知らなくても期待値を計算できるが、自己相関を生む。
2. HMC 法は作用の勾配で高次元空間を移動し、Metropolis 補正で数値積分誤差を除く。
3. 正規化フローは簡単な分布を可逆変換し、生成と密度評価を同時に可能にする。
4. 拡散モデルは加音過程の score を学び、Gauss ノイズから反復的に配位を生成する。
5. フローを MH 提案として使えば、学習器が不完全でも最終分布を厳密に保てる。
6. 残る課題は、体積・相関長・トポロジーに対するスケーリングを改善することである。

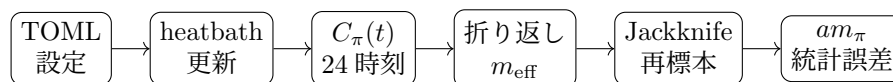
機械学習には効率のよい提案を作らせ、正しさは Metropolis–Hastings 補正で保証する。この分業が本講の結論である。

第3講

おまけ：LatticeQCD.jlでパイオン質量を測る

12 この実習で何を学ぶか

格子 QCD の計算は、パラメータファイルを実行して最後に一つの数字を読むだけの作業ではない。入力した作用、乱数の seed、更新履歴、測定間隔、解析から除いたゲージ配位、解析式までが一つの計算を構成する。本ハンズオンでは、その連鎖をパイオン質量の計算に絞って最後まで追う。



JuliaQCD は、格子 QCD の計算に必要な機能を Julia で実装するソフトウェア開発プロジェクトの総称である。Julia の実行時コンパイル (JIT)、多重ディスパッチ、MPI による並列化を利用し、手元の計算機での試作から並列計算機での実行まで、同じ Julia のプログラムを段階的に発展させることを目指している [42, 44]。本ハンズオンでは、その入口となる LatticeQCD.jl を使い、設定ファイルの読み込みからパイオン相関関数の測定までを実行する。

実行に必要な Julia 環境、設定、スクリプト、検証済みデータは、この講義リポジトリの `third_lecture/latticeqcd_tutorial` に同梱している。第 1・第 2 講で導入した格子作用、Monte Carlo 法、Dirac 演算子、相関関数、自己相関を、ここでは 1 本の計算として接続する。

範囲 ゲージ配位の生成、パイオン相関関数、有効質量、Jackknife 法、尺度変換を扱う。トポロジカルチャージや有限温度物理など、パイオン質量の計算に直接必要でない物理量は扱わない。

最終値は教育用の小体積クエンチ計算であり、実験値を再現する精密計算ではない。「なぜ値が出るか」と「なぜその値を物理値と呼べないか」の両方を説明できることが目標である。

13 パイオン質量を測るとはどういうことか

パイオンは、軽いクォークと反クォークから作られる擬スカラー中間子である。荷電パイオンを生成・消滅させる局所演算子を

$$P^+(x) = \bar{d}(x)\gamma_5 u(x), \quad P^-(x) = P^+(x)^\dagger \quad (176)$$

とする。フェルミオン作用 $S_f = \bar{q}D[U]q$ に対する Matthews-Salam (マッシュューズ-サラム) 公式は、Grassmann 変数のソース $\eta, \bar{\eta}$ を導入すると [43]

$$\int \mathcal{D}\bar{q}\mathcal{D}q e^{-\bar{q}Dq + \bar{\eta}q + \bar{q}\eta} = \det D e^{\bar{\eta}D^{-1}\eta} \quad (177)$$

と書ける。ソースで微分するとクォーク伝播関数

$$S_{ij}[U] = \langle q_i \bar{q}_j \rangle_F = (D^{-1}[U])_{ij} \quad (178)$$

が得られる。つまり必要なのは巨大な逆行列そのものではなく、選んだソースに対する D^{-1} の作用である。

色の添字 $a, b = 1, 2, 3$ とスピンの添字 $\alpha, \beta = 1, \dots, 4$ をまとめて $A = (a, \alpha)$ 、 $B = (b, \beta)$ と書く。原点に置く 12 個の点ソースを

$$\eta_B^{(A)}(y) = \delta_{y,0}\delta_{BA}, \quad \sum_{y,C} D_{BC}(x,y)\psi_C^{(A)}(y) = \eta_B^{(A)}(x) \quad (179)$$

を順に解けば

$$\psi_B^{(A)}(x) = S_{BA}(x, 0) \quad (180)$$

となり、1 回の求解で伝播関数の色・スピン成分に対応する 1 列が得られる。

CG (共役勾配法) は、残差 $r_k = \eta - D\psi_k$ を反復的に小さくして線形方程式を解く Krylov 部分空間法であり、 D^{-1} を明示的に保存しない。通常の CG は Hermitian 正定値行列に対する解法である。Wilson D は非 Hermitian 行列なので、この設定の `solve_DinvX!` は既定の BiCG を使い、

$$\|r_k\|^2 < \epsilon \quad \text{または} \quad k = \text{MaxCGstep} \quad (181)$$

まで反復する。ここでは $\epsilon = 10^{-19}$ 、`MaxCGstep`=3000 である。

最後に $D^\dagger = \gamma_5 D \gamma_5$ を使うと、運動量ゼロのパイオン相関関数は

$$C_\pi(t) = \left\langle \sum_{\mathbf{x}} \sum_{A,B} \left| \psi_B^{(A)}(\mathbf{x}, t) \right|^2 \right\rangle_U \xrightarrow[t \text{ large}]{} A_\pi \left(e^{-am_\pi t} + e^{-am_\pi(N_t-t)} \right) \quad (182)$$

と再構成できる。空間和が $\mathbf{p} = 0$ へ射影し、Euclid 時間方向の減衰率が am_π を与える。

12 回の反転計算は、12 個の逆行列を作るという意味ではない。各求解で D^{-1} の 1 列を得て、式 (182) の絶対値二乗をソース側とシンク側の全色・スピン成分、全空間点について足す。

14 Euclid 経路積分から Monte Carlo 法へ

QCD の期待値は、ゲージ場 U とクォーク場 $q, \bar{q}$ の経路積分で

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}U \mathcal{D}\bar{q} \mathcal{D}q e^{-S_g[U] - S_f[U, \bar{q}, q]} O \quad (183)$$

と表される [8]。Grassmann 積分を行うと、フェルミオン行列式とクォーク伝播関数が現れる。動的フェルミオンを含む full QCD（非クエンチ計算）では、ゲージ配位の確率分布にも海クォークの効果が入る。

本ハズオンでは、ゲージ配位を生成するときに

$$\det D[U] \longrightarrow 1 \quad (184)$$

とするクエンチ近似を使う。一方、測定時には各ゲージ場上で Wilson–Dirac 演算子を反転し、価クォークの伝播関数 $D^{-1}[U]$ を計算する。

N 個のゲージ配位 U_i に対する推定量は

$$\overline{C_\pi(t)} = \frac{1}{N} \sum_{i=1}^N C_\pi(t; U_i) \quad (185)$$

である。ただし U_i は独立な標本ではなく、Markov 連鎖上で続けて生成された標本である。自己相関を無視すると統計誤差を過小評価するおそれがあるため、block Jackknife 法（ブロック・ジャックナイフ法）も併用する。

ゲージ配位生成は「確率分布に従う標本を作る段階」、Dirac 演算子の反転は「各標本上でクォークの応答を測る段階」、Jackknife 法は「標本間の揺らぎを最終結果の誤差へ伝える段階」である。

15 他の格子 QCD コードと JuliaQCD の位置づけ

格子 QCD には、本格的な大規模計算で長年使われてきたソフトウェアが多数ある。各コードは、採用するフェルミオン作用、配位生成アルゴリズム、対応する計算機、入出力形式などにそれぞれ重点がある。したがって、言語や速度だけを見て一列に順位づけるものではなく、研究目的と計算環境に応じて選択する。JuliaQCD の論文では、MILC、CPS、openQCD、Bridge++、Grid/GPT、SIMULATEQCD、QUADA などが関連するソフトウェアとして紹介されている [42]。

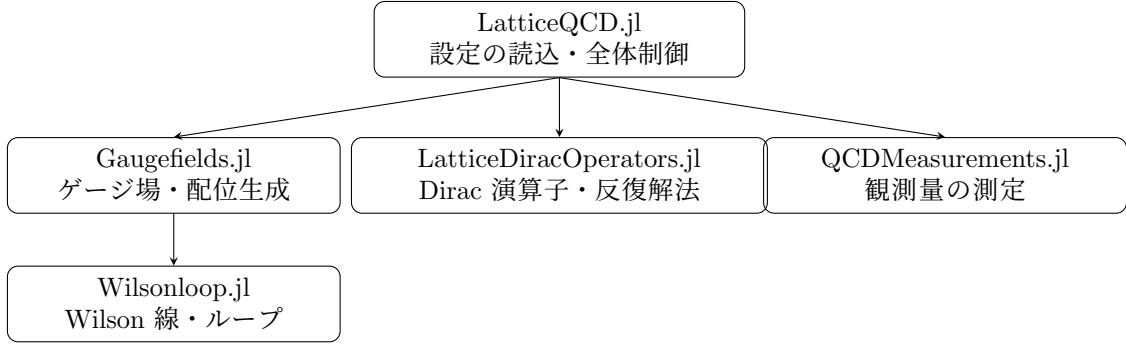
コード	主な特徴	パイオン質量計算との関係
MILC / CPS	C 言語を中心とする長期運用コード	大規模計算で生成された配位集合を扱う代表例
openQCD	改良 Wilson フェルミオン、HMC/SMD	Wilson 型フェルミオンを用いる高性能な配位生成
Bridge++	C++ で書かれた日本発の汎用基盤	実装方法や計算結果を照合する対象
Grid / GPT	SIMD、CPU/GPU、Python からの操作	現代的な並列計算と対話的な解析の接続
SIMULATEQCD / QUADA	CUDA、複数 GPU による高速化	Dirac 演算子などの GPU 計算を高速化
JuliaQCD	Julia による一貫した実装	数式、数値計算、解析の対応を追いやすい

JuliaQCD の特徴は、利用者が操作する上位部分と、計算時間の多くを占める数値計算部分を、原則として同じ Julia 言語で記述できる点にある。対話的に小さな格子で式やアルゴリズムを確認した後、同じ型と関数を使いながら MPI 並列計算へ進めることを意図している。ただし、Julia を使えば自動的にどの計算機でも最高速になるという意味ではない。大規模な精密計算では、実装の検証状況、利用可能な配位集合、対象となる GPU などの計算加速器、保守体制まで含めてコードを選ぶ必要がある。

本教材で JuliaQCD を使う理由は、単にコードが短いからではない。パイオン質量計算を、ゲージ場の生成、Dirac 方程式の求解、相関関数の構成、統計解析という段階に分け、各段階の数式と Julia の処理を対応させて読めるためである。

16 JuliaQCD を構成するパッケージ

「JuliaQCD」は一つの Julia パッケージの名前ではなく、共通の目的を持つ複数のパッケージからなるプロジェクト名である [42]。機能を分割することで、ゲージ場の更新法だけを調べる場合や、Dirac 方程式を解く新しい反復法を試す場合には、必要な部分を直接利用できる。一方、初学者が最初から内部の細かな関数やデータ型を組み合わせるのは難しい。そのため LatticeQCD.jl が上位の入口となり、TOML 形式の設定を読み、必要な下位パッケージを呼び出して、配位生成と測定の手順全体を管理する。



パイオン質量計算では、利用者の指示が次の順序で具体的な数値計算へ渡される。

1. LatticeQCD.jl が格子サイズ、 β 、更新回数、測定条件を読み、計算全体を初期化する。
2. Wilsonloop.jl のループ定義を用いて作用を組み立て、Gaugefields.jl が $U_\mu(x)$ を保持し、heatbath と overrelaxation で更新する。
3. QCDDMeasurements.jl がパイオン測定を開始し、LatticeDiracOperators.jl に Wilson-Dirac 演算子の構築と 12 本の色・スピン点源に対する線形方程式の求解を依頼する。
4. QCDDMeasurements.jl が得られた伝播関数を縮約し、各時刻の $C_\pi(t)$ を出力する。

利用者が直接実行するのは主に LatticeQCD.jl だが、実際の物理量は一つのパッケージだけで計算されるわけではない。この分担を理解すると、線形方程式の解法だけを変更する、別の測定を追加する、スミアリング（場の平滑化）を導入するといった拡張で、どのパッケージを調べればよいか判断できる。

17 再現可能な二つの実行経路

利用可能な時間と計算資源に応じて、二つの実行経路を選べる。

	短時間経路	完全実行経路
入力	付属の <code>Pion_correlator.txt</code>	TOML 設定ファイル
計算	読み込み、検査、統計解析	heatbath 更新、反転計算、統計解析
所要時間	数秒	Apple M1、1 スレッドで約 18 分
学ぶ点	相関関数からの質量抽出	配位生成を含む計算全体の再現
同じ部分	同じ <code>PionMassAnalysis.jl</code> を呼ぶ	

第三者が計算を再現できるように、リポジトリには次の情報を残す。

- `Project.toml` と `Manifest.toml`: パッケージの版と `git commit` (コミット)。
- `configs/*.toml`: 格子サイズ、 β 、乱数シード、 κ 、測定間隔。
- 未加工の相関関数: ゲージ配位ごとに N_t 個の値を並べたもの。
- 解析用 CSV: 平均、Jackknife 誤差、有効質量、解析区間の比較。
- 実行環境: OS、CPU、Julia、および各 JuliaQCD パッケージの版。

notebook の各セルで `Pkg.add` を繰り返すと、利用者全体に共通する環境が変化してしまう。リポジトリの最上位ディレクトリを Julia のプロジェクト環境とし、通常は `Pkg.instantiate()` だけを実行する。

18 Windows、macOS、Linux への Julia 導入

Julia 本体は `juliaup` で管理する。パッケージを固定するプロジェクト環境と、Julia 本体の版を選ぶ channel (系列) を分けて管理することで、更新のしやすさと計算の再現性を両立できる。

OS	導入コマンド
Windows	PowerShell で <code>winget install --name Julia --id 9NJNWW8PVKMN -e</code>
macOS / Linux	<code>curl -fsSL https://install.julialang.org sh</code>
全 OS	<code>juliaup add 1.12, juliaup default 1.12</code>
確認	<code>julia --version, juliaup status</code>

```
git clone <lecture-repository-url>
cd p2eak-lqcd-ml-lecture-2026/third_lecture/latticeqcd_tutorial
julia --project=. -e "using Pkg; Pkg.instantiate()"
julia --project=. test/runtests.jl
```

付属データを生成した実行では Julia 1.12.4 を使い、Project が許す Julia の版は 1.10、1.11、1.12 とした。Windows と Unix 系 OS で異なるパス区切りは Julia の `joinpath` が処理する。OS 間の移植性を保つには、shell に依存する `ls`、`tail`、`curl -O` ではなく、Julia の `readlines`、`Downloads.download`、`mkpath` を使う。

`julia` が見つからない場合は端末を開き直す。Windows では Store/winget 版、macOS/Linux では `juliaup` のインストール先が環境変数 `PATH` に含まれているかを確認する。

19 Project と Manifest を固定する

リポジトリの最上位に置く `Project.toml` には、この教材が直接利用する依存パッケージだけを記す。

```
1 using Pkg
2 Pkg.activate(".")
3 Pkg.instantiate()
4 Pkg.status()
```

直接依存として固定したパッケージは次の通りである。

```
JkTools v1.0.0-DEV
  github.com/akio-tomiya/JkTools#fa8f234
LatticeQCD v1.3.7
```

JkTools は General registry に登録されていないため、Manifest にはリポジトリの URL と commit (コミット) `fa8f234...` を記録した。LatticeQCD.jl の依存関係として選ばれた主要パッケージの版は、QCDMeasurements.jl 0.2.13、Gaugefields.jl 0.7.2、LatticeDiracOperators.jl 0.6.3、Wilsonloop.jl 0.1.5 である。

`Project.toml` は「許される直接依存」、`Manifest.toml` は「実際に選ばれた全依存」の記録である。Julia のマイナー版ごとに Manifest を分ける場合は、`Manifest-v1.11.toml` のように配置できる。

20 Google Colab: Julia を毎回入れ直さない

Colab ではランタイムの種類を Julia に切り替え、Colab が提供する Julia ランタイムを使う。notebook 内で Julia 本体と IJulia を毎回入れ直す方法では、大きなファイルの取得と kernel の再起動が毎回必要になるため、起動時間が長くなる。

最初のセルでは、教材のプロジェクト環境を有効化し、必要な依存関係を復元する。

```
1 using Pkg
2 Pkg.activate("/content/p2eak-lqcd-ml-lecture-2026/third_lecture/latticeqcd_tutorial"
3 )
4 @time Pkg.instantiate()
5 Pkg.status()
```

環境の復元が終わったら、付属の相関関数を使う短時間経路を実行する。

```
1 cd("/content/p2eak-lqcd-ml-lecture-2026/third_lecture/latticeqcd_tutorial")
2 include("scripts/analyze_pion.jl")
```

正常に終了すると、最後に次の結果が表示される。

```
pion correlators read: 25
pion correlators analyzed: 21
am_pi = 1.331004 +/- 0.048303 (block JK error, block=2)
m_pi = 1539.0 +/- 55.9 MeV (statistical only)
```

ここまでが授業時間内に行う解析である。ゲージ配位の生成から実行する場合は、第 16 節に示すシミュレーション用スクリプトを使う。

21 使用する物理・数値パラメータ

項目	値	記号	意味
格子サイズ	$4^3 \times 24$	$N_s^3 \times N_t$	小さな空間体積、長い時間方向
ゲージ結合	5.7	$\beta = 6/g_0^2$	Wilson plaquette ゲージ作用
更新法	heatbath + 2 OR	–	クエンチ近似でのゲージ場更新
初期配位	cold	$U_\mu = 1$	熱化が必要な単位配位
乱数シード	20260718	–	擬似乱数列を再現するため
価フェルミオン	Wilson	–	パイオン伝播関数の測定に使用
hopping parameter	0.141139	κ	クォーク質量を制御
測定間隔	5 updates ごと	–	120 回の更新から 24 回測定

尺度の補間式から $a \simeq 0.171 \text{ fm}$ と置くと

$$L_s = N_s a \simeq 0.68 \text{ fm}, \quad L_t = N_t a \simeq 4.10 \text{ fm} \quad (186)$$

となる。時間方向は相関関数の減衰を見るため長くし、空間方向は手元の計算機で短時間に反転計算を終えるため、物理計算としては極端に小さくしている。

このパラメータはアルゴリズムの実習用である。有限体積効果を抑える経験則 $m_\pi L_s \gtrsim 4$ を満たさず、連続極限への外挿、海クォーク、複数の格子間隔も含んでいない。

22 Heatbath 更新、overrelaxation、熱化

ゲージリンク $U_\mu(x) \in SU(3)$ は Wilson plaquette 作用

$$S_g[U] = \beta \sum_{x, \mu < \nu} \left(1 - \frac{1}{3} \text{Re tr} U_{\mu\nu}(x) \right) \quad (187)$$

に従う。heatbath 法は、周囲のリンクを固定したときの条件付き確率分布に従ってリンクを更新する。overrelaxation は、作用を保つ方向へゲージ場を大きく動かし、自己相関の低減を狙う。格子 Monte Carlo 法の基礎は Wilson と Creutz の仕事に遡る [8, 15]。

cold start はすべてのリンクを単位行列にした初期状態であり、平衡分布から取り出した配位ではない。付属の未加工ファイルには、LatticeQCD.jl が準備確認として出力した trajectory 番号 0 の相関関数 1 本と、trajectory 番号 5, 10, ..., 120 の相関関数 24 本があり、合計 25 行を含む。解析では trajectory 番号 0, 5, 10, 15 を除き、20 から 120 までの 21 本を使う。

区間	扱い	理由
0, 5, 10, 15	解析から除外	cold start の影響を避ける
20–120	解析に使用	5 updates 間隔、相関関数 21 本
block size 2	誤差を確認	10 blocks、余る先頭 1 本を除く

Nthermalization=20 という設定値だけで判断せず、出力ファイルの実際の行数と trajectory 番号の対応を確認する。このパイオン相関関数ファイルには trajectory 番号の列がないため、測定順序を設定ファイルと一緒に保存する。

23 Wilson フェルミオンと hopping parameter

Wilson–Dirac 演算子の一例は

$$D_{xy} = \delta_{xy} - \kappa \sum_{\mu=1}^4 \left[(1 - \gamma_{\mu}) U_{\mu}(x) \delta_{x+\hat{\mu},y} + (1 + \gamma_{\mu}) U_{\mu}^{\dagger}(x - \hat{\mu}) \delta_{x-\hat{\mu},y} \right] \quad (188)$$

である。Wilson 項は余分なフェルミオンモードである doubler を重くする一方、有限の格子間隔ではカイラル対称性を明示的に破る。裸のクォーク質量は概ね

$$am_0 = \frac{1}{2\kappa} - 4r \quad (189)$$

と関係する。質量ゼロに対応する臨界値 κ_c は相互作用によって自由理論の値からずれる。ここでは $r = 1$ 、 $\kappa = 0.141139$ とし、clover 改良は用いない。

点源からの伝播関数 $S = D^{-1}\eta$ を作るには、ソース点の色 3 成分とスピン 4 成分を組み合わせ、合計 12 本の線形方程式を解く。実行ログの Inversion 1/12 から 12/12 はこの進捗である。

ゲージ配位の生成はクエンチ近似だが、パイオン測定では Wilson 型の価クォーク伝播関数を使う。「フェルミオンがまったく無い」のではなく、フェルミオン行列式をゲージ配位の確率重みに入れていないという意味である。

24 パイオン演算子と伝播関数の縮約

Wick 縮約と γ_5 hermiticity $S(0, x) = \gamma_5 S(x, 0)^\dagger \gamma_5$ を使う。演算子の位相を式 (182) の正の規約に選ぶと、isovector パイオンの connected 相関関数は

$$C_\pi(t) = \sum_x \left\langle \text{Tr}_{c,s} \left[S(x, 0) S(x, 0)^\dagger \right] \right\rangle_U \quad (190)$$

となる。したがって各ゲージ配位上では、12 本の点ソースに対する解の絶対値二乗を、シンク側の全色・スピン成分と全空間点について加えればよい。

QCDMeasurements.jl のパイオン測定は次の順に進む。

1. 原点に色・スピン成分を指定した点ソースを置く。
2. Wilson-Dirac 演算子を 12 個のソースについて反転する。
3. シンク時刻 t ごとに、全空間点と内部添字を縮約する。
4. $t = 0, \dots, 23$ の 24 個の値を 1 行に書き、末尾に `#pioncorrelator` を付ける。

点ソースは実装が単純である。一方、基底状態との重なりを改善するには、ソースとシンクにスミアリングを施す方法が有効である。スミアリングを変えると plateau が現れ始める時刻も変わり得るため、質量を平均する時間区間を選び直す必要がある。

1 行は 1 時刻の値ではなく、1 ゲージ配位に対する 24 時刻すべての値である。行列として読んだときの形状は $(N_{\text{config}}, N_t) = (25, 24)$ となる。

25 周期格子の相関関数と cosh

転送行列によるスペクトル分解から、運動量ゼロの相関関数は

$$C_\pi(t) = \sum_n |Z_n|^2 \left(e^{-E_n t} + e^{-E_n(N_t-t)} \right) \quad (191)$$

となる。十分大きい t でパイオンの基底状態が支配的になれば

$$C_\pi(t) \simeq A \left(e^{-am_\pi t} + e^{-am_\pi(N_t-t)} \right) = 2Ae^{-am_\pi N_t/2} \cosh \left[am_\pi \left(t - \frac{N_t}{2} \right) \right]. \quad (192)$$

まず時間反転を平均して

$$C_\pi^{\text{fold}}(t) = \frac{C_\pi(t) + C_\pi(N_t - t)}{2}, \quad 1 \leq t < N_t/2 \quad (193)$$

とする。次に cosh 型の有効質量

$$am_{\text{eff}}(t) = \text{arccosh} \left[\frac{C_\pi(t-1) + C_\pi(t+1)}{2C_\pi(t)} \right] \quad (194)$$

を計算する。

$-\log[C(t+1)/C(t)]$ は、時間境界を反対向きに回り込む項を無視するため、 $N_t/2$ に近づくほど不適切になる。この対数比は指数減衰の直観的な確認にだけ使い、最終値には cosh 形式を使う。

26 実行に用いる TOML 設定ファイル

以下が実際に使用した設定である。物理パラメータと出力先をファイルに保存することで、設定作成 wizard の対話履歴に依存せず、同じ条件を再利用できる。

```

1  ["HMC related"]
2  "Δτ" = 0.05
3  MDsteps = 20
4
5  ["Physical setting(fermions)"]
6  Dirac_operator = "nothing"
7
8  ["Physical setting"]
9  L = [4, 4, 4, 24]
10 "β" = 5.7
11 update_method = "Heatbath"
12 initial = "cold"
13 useOR = true
14 numOR = 2
15 Nthermalization = 20
16 Nsteps = 120
17
18 ["System Control"]
19 logfile = "pion_mass_4x4x4x24.log"
20 verboselevel = 2
21 randomseed = 20260718
22 measurement_dir = "pion_mass_4x4x4x24"
23 measurement_basedir = "./results/verified_run/measurements"
24 log_dir = "./results/verified_run/logs"
25 hasgradientflow = false
26
27 ["Measurement set".measurement_methods.Pion_correlator]
28 methodname = "Pion_correlator"
29 measure_every = 5
30 eps = 1.0e-19
31 MaxCGstep = 3000
32 printvalues = true
33
34   ["Measurement set".measurement_methods.Pion_correlator.fermion_parameters]
35   Dirac_operator = "Wilson"
36   hop = 0.141139
37   r = 1.0
38   hasclover = false
39
40 [gradientflow_measurements]
41 hasgradientflow = false
42
43   [gradientflow_measurements.measurements_for_flow]
```

特に randomseed、measure_every、Wilson の hop を省略しない。Dirac_operator="nothing" はゲージ場の更新がクエンチ近似であることを表す。一方、入れ子の表にある Dirac_operator="Wilson" は、測定時に Wilson 型の価クォークを使う指定であり、両者の役割は異なる。

27 シミュレーションを実行し、ログを読む

リポジトリの最上位ディレクトリで次を実行する。

```
julia --project=. scripts/run_simulation.jl \
  configs/pion_mass_4x4x4x24.toml
```

実行出力の抜粋は次の通りである。

```
$ julia --project=. scripts/run_simulation.jl configs/pion_mass_4x4x4x24.toml
tutorial_julia_version = 1.12.4
LatticeQCD 1.3.7 / Gaugefields 0.7.2
Heatbath update with even-odd method. Now beta = 5.7
Hadron spectrum: Inversion 1/12 is done
...
Hadron spectrum: Inversion 12/12 is done
Acceptance 120/120 : 100.0 percent
Total Elapsed time 1045.67093975 [s]
tutorial_simulation_wall_seconds = 1090.978
```

測定項目	Apple M1、1 スレッド
シミュレーションの実時間	1090.98 s
パイオン測定	24 回 + 初期確認 1 回
1 回の測定に含まれる Dirac 反転	12 回

Acceptance 120/120 は heatbath 更新が 120 回完了したことを示し、HMC の Metropolis 受理率とは異なる。ここで確認すべき点は、error を出さずに 120 回の更新とすべての測定が完了し、未加工ファイルに 25 行が出力されたことである。

28 未加工の相関関数を厳密に読む

読み込み関数は、各行の comment より前だけを数値に変換する。さらに、各行の数値が N_t 個、または将来 trajectory 番号の列が追加された場合の $N_t + 1$ 個であることを検査する。

```

1 for (line_number, rawline) in enumerate(eachline(filename))
2     line = strip(first(split(rawline, '#'; limit=2)))
3     isempty(line) && continue
4     parsed = parse.(Float64, split(line))
5
6     if length(parsed) == nt
7         push!(rows, parsed)
8     elseif length(parsed) == nt + 1
9         push!(trajectories, Int(parsed[1]))
10        push!(rows, parsed[2:end])
11    else
12        error("line $line_number: unexpected number of fields")
13    end
14 end
15 all(isfinite, reduce(vcat, permutedims.(rows))) || error("non-finite data")

```

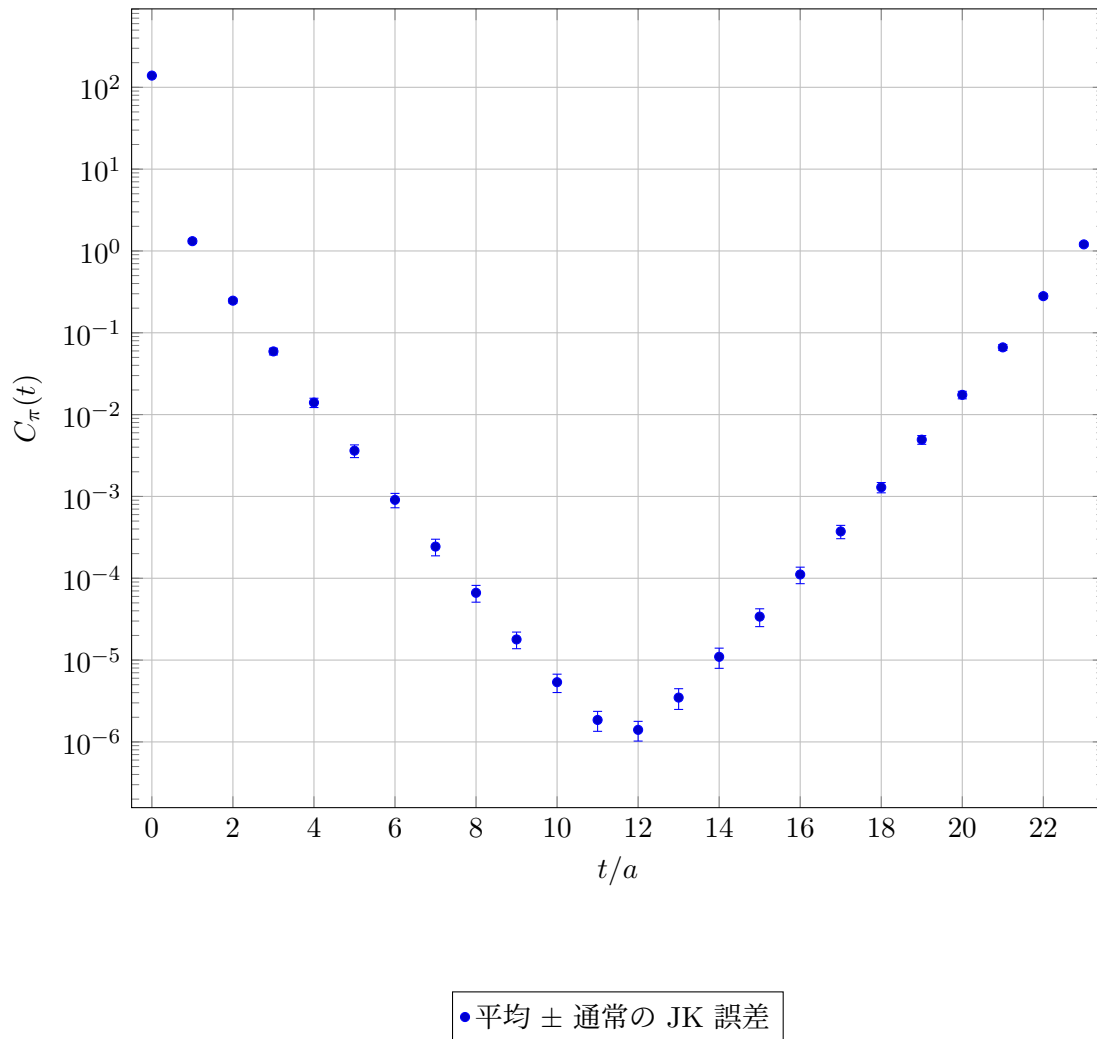
付属データを検査した結果は

$$N_{\text{raw}} = 25, \quad N_t = 24, \quad N_{\text{discard}} = 4, \quad N_{\text{analyze}} = 21 \quad (195)$$

となった。trajectory 番号の列がないため、最初の確認測定を 0、その後を 5 刻みとして設定ファイルから対応を復元する。

29 平均、時間反転、相関関数の形

trajectory 番号 20–120 の 21 本を平均し、時間の正方向と負方向を折り返して平均した。図には $t = 0 \dots 23$ の全範囲について、平均値と通常の Jackknife 誤差を示す。

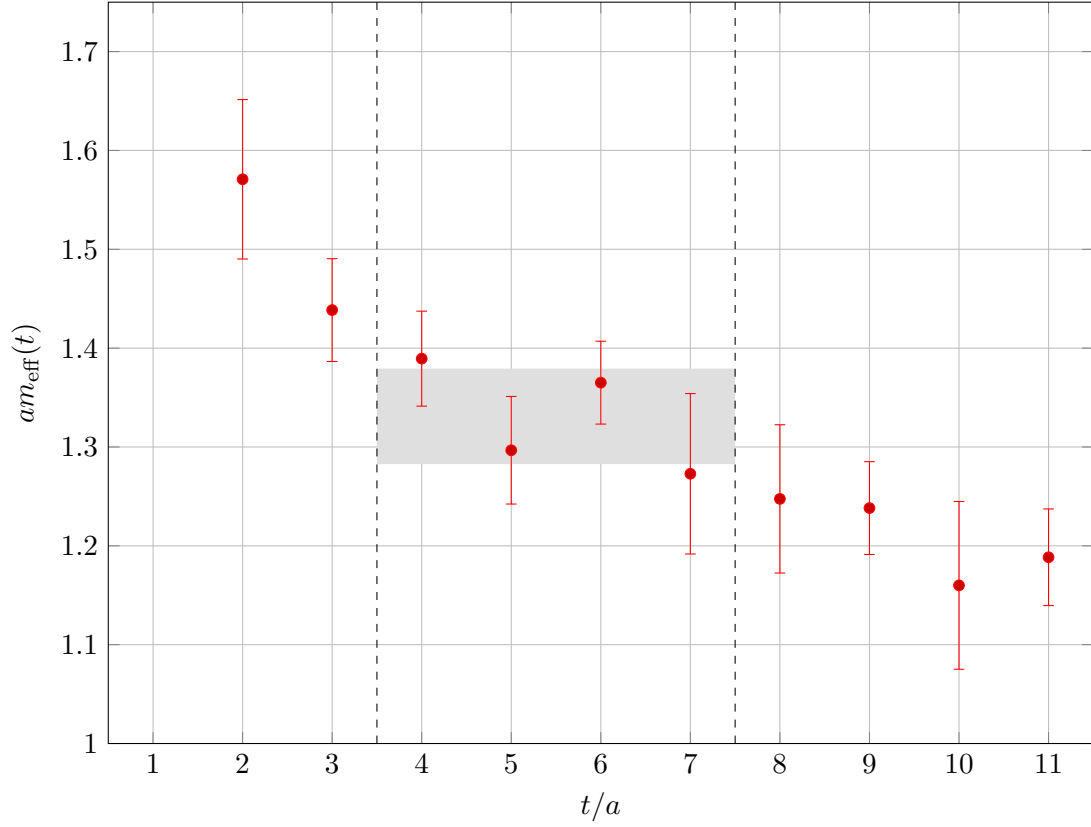


$t = 0$ のソース近傍から急速に減衰し、 $t = 12$ で最小となった後、再び増加する。これは周期的な時間境界を反対向きに回り込む伝播のためであり、計算が壊れた徴候ではない。

図を作る前に平均値だけへ縮約せず、ゲージ配位を表す軸を解析の最後まで残す。時刻ごとの平均しか保存しないと Jackknife 再標本を作れないため、元の標本を捨ててはいけない。

30 cosh 型有効質量と plateau

時間反転で折り返した相関関数から、cosh 型有効質量を計算した。帯は採用値 1.3310 ± 0.0483 、縦の破線は plateau とみなして平均する区間 $t = 4 \dots 7$ を示す。



時間区間	am_{π}	block-JK 誤差
3 : 6	1.3724	0.0363
4 : 7	1.3310	0.0483
5 : 8	1.2955	0.0522
6 : 9	1.2809	0.0469
4 : 9	1.3016	0.0448

時間区間を変えたときの値の変化は、統計誤差と同程度か、それより大きい。採用した区間は一意に決まる精密なフィットの結果ではなく、励起状態の混入と統計的な雑音の両方を見ながら選んだ範囲である。

31 Jackknife 法と JkTools: 非線形解析を丸ごと再計算

N 個の標本から 1 個ずつ除いて計算した推定値を $\theta_{(i)}$ とすると

$$\bar{\theta}_{\text{JK}} = \frac{1}{N} \sum_i \theta_{(i)}, \quad \sigma_{\text{JK}} = \sqrt{\frac{N-1}{N} \sum_i (\theta_{(i)} - \bar{\theta}_{\text{JK}})^2} \quad (196)$$

である [45, 46]。 θ は単純平均である必要はない。ここでは、「相関関数の平均 $\rightarrow$ 時間反転による折り返し $\rightarrow$ arccosh $\rightarrow$ 時間区間内の平均」という非線形な解析手順全体を、標本を除くたびに最初から計算し直す。

```

1 subsets = jk_index_set(1:nconfig; block=2)
2 replicas = [
3     plateau_mass(values[index, :]; tmin=4, tmax=7)
4     for index in subsets
5 ]
6 nrep = length(replicas)
7 err = sqrt((nrep - 1) / nrep *
8           sum((replicas .- mean(replicas)).^2))

```

block size	中心値	JK 誤差	再標本数 / 使用標本数
1	1.3310	0.0328	21 / 21
2	1.3382	0.0483	10 / 20
3	1.3310	0.0576	7 / 21
4	1.3382	0.0450	5 / 20

JkTools の block 用関数は、標本数が block size で割り切れない場合、先頭の余りを除いて同じ大きさの block を作る。block size 2 では 21 本中先頭 1 本を除き、10 blocks となる。block size 4 では再標本数が 5 まで減るため、その誤差推定値を過信してはいけない。

32 最終結果、限界、次の演習

上記のパラメータと解析手順で得た結果は

$$am_\pi = 1.3310 \pm 0.0483 \quad (\text{block Jackknife, block}=2) \quad (197)$$

$$m_\pi = 1539 \pm 56 \text{ MeV} \quad (\text{統計誤差のみ}) \quad (198)$$

格子間隔は Guagnelli–Sommer–Wittig によるクエンチ近似の補間式 [47]

$$\ln(a/r_0) = -1.6805 - 1.7139(\beta - 6) + 0.8155(\beta - 6)^2 - 0.6667(\beta - 6)^3 \quad (199)$$

に $r_0 = 0.5 \text{ fm}$ 、 $\beta = 5.7$ を入れ、 $a = 0.17066 \text{ fm}$ とした。尺度設定に由来する不定性は、上の 56 MeV には含めていない。

結論の読み方: この計算では、統計誤差よりも、解析する時間区間の選び方、有限体積、クエンチ近似、離散化、尺度設定に由来する系統的不定性の方が大きい。したがってこの値は、現実のパイオン質量に対する精密予言ではなく、相関関数から質量を抽出する一連の手順を再現した結果である。

演習:

1. 時間区間を $3:6, 5:8, 6:9$ に変え、中心値の変化と統計誤差の大小を説明する。
2. block size を $1-4$ で変え、再標本数が減ると誤差推定が不安定になる理由を述べる。
3. κ を一つだけ変え、同じ乱数シードと統計量で am_π の変化を見る。
4. 新しい Colab ランタイムで、依存関係の復元、最初のパッケージ読み込み、シミュレーションを別々に計時する。

再実行時に呼ぶスクリプトは `run_simulation.jl` と `analyze_pion.jl` の二つである。未加工データ、解析用 CSV、結果の要約、実行環境の記録、この TeX ノートを一緒に保存し、一つの成果物とする。

補足資料

A1 くりこみ群変換で何が不変になるか

すべての期待値の数値を、そのまま同時に不変にできるという意味ではない。完全な Wilsonian くりこみ群では、カットオフ変更によって生成される、対称性が許すすべての演算子と外場 (source) 項を有効作用に残し、場と演算子の規格化も同時に変換する。この無限次元の理論空間で変換を厳密に行えば、低エネルギーの生成汎関数、したがって物理的な相関関数の内容を保つことができる [12]。

一方、実際の計算で作用を有限個の演算子に切り詰めると、厳密に固定できるのは通常、くりこみ条件に選んだ有限個の物理量だけである。それ以外の物理量には、捨てた非重要演算子 (irrelevant operator) に由来する有限カットオフ効果が残り、連続極限で一定値へ近づく。また、場の相関関数そのものは正準次元と異常次元に従って変換するため、その数値が文字どおり不変とは限らない。不変なのは、適切な場・演算子の変換を含めて比較した物理的予言である。

A2 虚時間相関関数と質量

基底状態 $|0\rangle$ に対し、虚時間 Heisenberg 演算子を

$$O(\tau) = e^{H\tau} O(0) e^{-H\tau} \quad (200)$$

と定義する。2 点相関関数に完全系 $\sum_n |n\rangle\langle n| = 1$ を挿入すると

$$C(\tau) = \langle 0 | O(\tau) O^\dagger(0) | 0 \rangle \quad (201)$$

$$= \sum_n |\langle 0 | O | n \rangle|^2 e^{-(E_n - E_0)\tau}. \quad (202)$$

大きな τ では、 O が結合する最も低い励起状態が支配する。真空エネルギーを差し引いた相対論的な 1 粒子の励起エネルギーは

$$E(\vec{p}) - E_0 = \sqrt{\vec{p}^2 + m^2} \quad (203)$$

である。そこで演算子を空間方向に和して $\vec{p} = \vec{0}$ へ射影すると、 $E(\vec{0}) - E_0 = m$ となる。このとき

$$C(\tau) \sim A e^{-m\tau} \quad (204)$$

の減衰率から、そのチャンネルで最も軽い粒子の質量 m が得られる。虚時間で質量を計算できるのは、同じ時間非依存ハミルトニアン固有値が指数減衰率として直接現れるからである。

A3 Euclid 相関関数と再構成定理

虚時間を含む Euclid 相関関数は経路積分から計算できるが、たとえ閉じた形で求まっても、形式的に $\tau \rightarrow it$ と代入するだけで Minkowski 相関関数が得られるとは限らない。必要な領域での解析性や、分布としての境界値の存在が自動的に保証されるわけではないからである。Osterwalder–Schrader 再構成定理は、Euclid 相関関数 (Schwinger 関数) が Euclid 共変性、対称性、反射正值性、適切な正則性やクラスター性などの OS 公理を満たせば、Wightman 公理を満たす Minkowski 量子場理論を再構成

できることを述べる [1]。一方、Wightman 再構成定理は、適切な Wightman 関数から Hilbert 空間、真空および場の演算子を再構成する [2]。したがって、虚時間から実時間への移行を正当化するのは単なる変数置換ではなく、相関関数がこれらの再構成定理の仮定を満たすことである。

実時間への解析接続を行わず、Euclid 形式から直接対応づけやすいのは、Minkowski 時空で空間的 (spacelike) に離れた相関である。計量を $(+, -, -, -)$ とすれば、

$$\Delta s_M^2 = (\Delta t)^2 - |\Delta \mathbf{x}|^2 \quad (205)$$

であり、 $\Delta s_M^2 < 0$ なら Lorentz 変換によって $\Delta t' = 0$ の系を選べるため、 $r = \sqrt{-\Delta s_M^2}$ として、スカラー演算子の 2 点関数は模式的に

$$G_M(\Delta t, \Delta \mathbf{x})|_{\Delta s_M^2 < 0} = G_M(0, \mathbf{r}) \longleftrightarrow G_E(0, \mathbf{r}) \quad (206)$$

と等時刻の Euclid 相関関数へ対応する。一方、時間的分離 $\Delta s_M^2 > 0$ は Lorentz 変換で等時刻にできず、光的分離 $\Delta s_M^2 = 0$ は空間的領域と時間的領域の境界で、相関関数が特異な分布になることもある。それでも格子 QCD から静的クォーク間ポテンシャル、ハドロンの行列要素、および形状因子を計算できるのは、これらが実時間相関関数全体ではなく、Euclid 相関関数のスペクトル分解に現れる指数減衰率とその係数から読み取れるためである。静的ポテンシャル $V(R)$ は長方形 Wilson loop の大きな Euclid 時間 T における $\langle W(R, T) \rangle \sim e^{-V(R)T}$ から、行列要素は適切な生成演算子と消滅演算子の間に電流を挿入した 3 点相関関数と 2 点相関関数の比から求まる。また、初期状態と終状態に異なる運動量を与えれば、その行列要素の運動量移行依存性として形状因子が得られる。いずれも十分な Euclid 時間間隔で目的の状態を分離することが基本であり、実時間相関関数全体の逆 Laplace 変換を必要としない。一方、実時間スペクトルの数値的再構成は別の難しさを持ち、付録 A23 で説明する。

A4 差分演算子の古典連続極限

テイラー展開

$$\phi(x + a\hat{\mu}) = \phi(x) + a\partial_\mu\phi(x) + \frac{a^2}{2}\partial_\mu^2\phi(x) + O(a^3) \quad (207)$$

から

$$\frac{\phi(x + a\hat{\mu}) - \phi(x)}{a} = \partial_\mu\phi(x) + O(a) \quad (208)$$

である。対称差分なら誤差は $O(a^2)$ になる。これは古典作用の連続極限である。

しかし量子論では、単に $a \rightarrow 0$ と置くだけでは足りない。格子間隔を変えると短距離揺らぎの取り込み方も変わるため、裸のパラメータ m_0, λ_0 を a とともに調整し、質量比などの物理量を一定に保つ必要がある。

式 (31) の差分項に上のテイラー展開を代入し、場と作用を次元付きの表式へ戻して $\sum_n a^4 \rightarrow \int d^4x$ と書き直すと、格子作用は一般に

$$\begin{aligned} S_{\text{lat}} &= S_{\text{cont}} + \frac{c_5}{\Lambda} \int d^4x \mathcal{O}_5(x) + \frac{c_6}{\Lambda^2} \int d^4x \mathcal{O}_6(x) + \cdots \\ &= S_{\text{cont}} + \sum_{d_i > 4} \frac{c_i}{\Lambda^{d_i-4}} \int d^4x \mathcal{O}_i(x), \quad \Lambda = \frac{1}{a} \end{aligned} \quad (209)$$

と表せる。 $\mathcal{O}_i$ は質量次元 $d_i > 4$ の局所演算子であり、連続理論に加わる非関連項を表す。したがって格子作用は、連続作用そのものではなく、紫外カットオフ Λ の逆べきで抑えられた非関連項を含む作用と見なせる。これが Symanzik の有効作用による見方である。

式 (30) について具体的に計算してみよう。前進差分へ上の Taylor 展開を代入すると

$$\frac{\phi(x + a\hat{\mu}) - \phi(x)}{a} = \partial_\mu \phi + \frac{a}{2} \partial_\mu^2 \phi + \frac{a^2}{6} \partial_\mu^3 \phi + O(a^3)$$

となる。これを差分項の 2 乗へ代入し、周期境界条件を仮定して部分積分すると、 $O(a)$ 項は全微分として消える。従ってツリーレベルでは

$$\begin{aligned} S_{\text{lat}} &= \int d^4x \left[\frac{1}{2} \sum_\mu (\partial_\mu \phi)^2 + \frac{m_0^2}{2} \phi^2 + \frac{\lambda_0}{4!} \phi^4 - \frac{a^2}{24} \sum_\mu (\partial_\mu^2 \phi)^2 + O(a^4) \right] \\ &= S_{\text{cont}} - \frac{1}{24\Lambda^2} \int d^4x \sum_\mu (\partial_\mu^2 \phi)^2 + O(\Lambda^{-4}), \quad \Lambda = \frac{1}{a}. \end{aligned} \quad (210)$$

ここでは式 (209) の具体例として、 $\mathcal{O}_6 = \sum_\mu (\partial_\mu^2 \phi)^2$ と $c_6 = -1/24$ が現れている。量子補正まで含めると、同じ対称性をもつ他の高次元演算子との混合も生じる。

なお、実際にどの項が現れるかは格子作用の対称性で決まる。式 (31) のような標準的な反射対称のスカラー作用では、 $1/\Lambda$ に比例する項は全微分または対称性によって消え、主要な離散化誤差は $1/\Lambda^2 = a^2$ から始まる。

A5 積分測度、自由度とエリツツァーの定理

ゲージ場の分配関数は

$$Z = \int \mathcal{D}U e^{-S_g[U]}, \quad \mathcal{D}U = \prod_{n,\mu} dU_\mu(n) \quad (211)$$

である。 dU は $\text{SU}(3)$ 上の Haar 測度で、群の左作用・右作用に対して不変である。 $N_s^3 \times N_t$ 格子には $4N_s^3 N_t$ 本のリンクがあり、各リンクは $\text{SU}(3)$ の 8 自由度を持つ。例えば $32^3 \times 64$ なら約 6700 万の実自由度である。

格子上的の局所ゲージ変換を

$$U_\mu(n) \mapsto U_\mu^\Omega(n) = \Omega(n) U_\mu(n) \Omega^\dagger(n + \hat{\mu}), \quad \Omega(n) \in \text{SU}(3) \quad (212)$$

とする。右辺の 2 つの行列は、リンクの始点と終点で色空間の基底を回転する。Haar 測度は左移動と右移動に対して不変なので、各リンクについて

$$dU_\mu^\Omega(n) = d[\Omega(n) U_\mu(n) \Omega^\dagger(n + \hat{\mu})] = dU_\mu(n) \quad (213)$$

である。従って全リンクの積分測度も

$$\mathcal{D}U^\Omega = \prod_{n,\mu} dU_\mu^\Omega(n) = \prod_{n,\mu} dU_\mu(n) = \mathcal{D}U \quad (214)$$

を満たす。

Wilson 作用では、プラケットを一周する積の内部に現れる $\Omega^\dagger(n) \Omega(n)$ が隣り合うリンク間で相殺する。始点には共役変換だけが残るが、トレースの巡回性から

$$\text{Tr} U_{\mu\nu}^\Omega(n) = \text{Tr} [\Omega(n) U_{\mu\nu}(n) \Omega^\dagger(n)] = \text{Tr} U_{\mu\nu}(n) \quad (215)$$

であり、 $S_g[U^\Omega] = S_g[U]$ となる。

演算子 $O[U]$ のゲージ変換を $O^\Omega[U] = O[U^\Omega]$ と書くと、その期待値は

$$\begin{aligned} \langle O^\Omega \rangle &= \frac{1}{Z} \int \mathcal{D}U e^{-S_g[U]} O[U^\Omega] \\ &= \frac{1}{Z} \int \mathcal{D}V e^{-S_g[V]} O[V] = \langle O \rangle \end{aligned} \quad (216)$$

である。第2行では $V = U^\Omega$ と変数変換し、Haar 測度と作用の不変性を用いた。すなわち、積分変数の各リンクに任意の固定したゲージ変換行列を挿入しても、期待値は変わらない。

次に1つの格子点 n_0 だけでゲージ変換する行列を $h \in \text{SU}(3)$ とし、他の格子点では単位行列とする。この局所変換による群平均を

$$(\mathcal{P}_{n_0} O)[U] = \int_{\text{SU}(3)} dh O[U^{h_{n_0}}] \quad (217)$$

と定義する。規格化した Haar 測度について $\int dh = 1$ である。式 (216) は各 h について成り立つので、

$$\langle O \rangle = \int dh \langle O^{h_{n_0}} \rangle = \langle \mathcal{P}_{n_0} O \rangle \quad (218)$$

を得る。 $\mathcal{P}_{n_0}$ は、演算子から点 n_0 におけるゲージ一重項成分だけを取り出す射影である。

例えば O_a が点 n_0 で非自明な既約表現 $D(h)$ に従って $O_a \mapsto D_{ab}(h)O_b$ と変換するなら、表現の直交関係から

$$\int dh D_{ab}(h) = 0, \quad \mathcal{P}_{n_0} O_a = 0 \quad (219)$$

である。従ってゲージ依存な量の期待値は

$$\boxed{\langle O_a \rangle = 0} \quad (220)$$

となる。

エリツァーの定理

ゲージ固定を行わないコンパクト格子ゲージ理論では、局所ゲージ群の非自明な表現に属する局所演算子、より正確にはゲージ軌道上の群平均が0となる局所演算子の期待値は0である。従って局所ゲージ対称性は自発的に破れない [10]。^a

^a局所ゲージ対称性は、異なる物理状態を結ぶ通常の対称性ではなく、同じ物理状態を記述する際の冗長性である。Gauss 則を満たす物理的 Hilbert 空間では局所ゲージ変換は自明に作用するため、その多重項は Hamiltonian のスペクトルには現れない。ここで「自発的に破れない」とは、ゲージ変換に非自明に応答する局所秩序変数の、ゲージ固定を行わない期待値が、熱力学極限でも0以外の値を取れないという意味である。境界で非自明な変換や、ゲージ固定後に残る大域対称性は別に考える必要がある。

例えば単独リンクの始点 n_0 だけを変換すれば

$$\mathcal{P}_{n_0} U_\mu(n_0) = \int dh h U_\mu(n_0) = 0 \quad (221)$$

であり、開いた Wilson line にも同じ議論が適用できる。一方、閉じた Wilson loop では経路上の変換行列が隣り合うリンク間で相殺し、始点に残る共役変換もトレースによって消えるため、非零の期待値を持ち得る。ゲージ固定後にゲージ依存な量が非零になっても、それはエリツァーの定理と矛盾しない。

以上はゲージ対称な作用を用いた零外場での議論である。自発的対称性の破れを調べるには小さな対称性破り外場を加え、熱力学極限を取った後に外場を0へ戻す必要がある。局所ゲージ変換が動かすのは n_0 近傍の有限個の自由度だけなので、この順序で極限を取っても局所的な非一重項秩序変数は非零にならない。この主張を外場に対する評価まで含めて示したものがエリツァーの定理である。ゲージ固定を加えた理論や残留する大域対称性については、この結論をそのまま適用できない。

A6 代表的なゲージ不変観測量

ゲージ不変な観測量の基本は閉じた Wilson loop

$$W(C) = \frac{1}{N_c} \text{Tr} \prod_{\ell \in C} U_\ell \quad (222)$$

である。積は、向きを定めた閉路 C に沿ってリンクを順番に掛ける。経路を逆向きにたどるリンクには $U_\ell^\dagger$ を用いる。

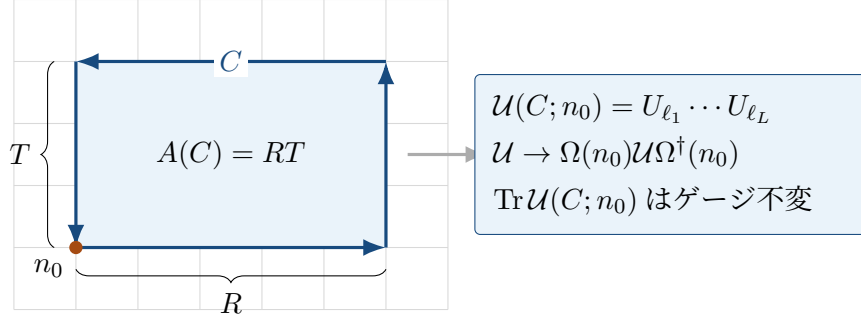


図 17: Wilson loop の模式図。閉路上の中間点では隣接するリンクの $\Omega^\dagger(n)\Omega(n)$ が相殺され、閉路積には始点 n_0 での共役変換だけが残る。そのトレースを取った式 (222) はゲージ不変である。

長方形 $R \times T$ に対し

$$\langle W(R, T) \rangle \sim e^{-V(R)T} \quad (223)$$

から静的クォーク間ポテンシャル $V(R)$ を抽出できる。最小の Wilson loop である平均プラケット

$$P = \frac{1}{6V} \sum_{n, \mu < \nu} \frac{1}{N_c} \text{Re Tr } U_{\mu\nu}(n) \quad (224)$$

は、熱化の確認やコードの検算に最初に用いられる観測量である。有限温度における Polyakov loop と中心対称性については、付録 A37 で説明する。ハドロン質量を求める場合は、クォーク伝播関数を組み合わせた中間子・バリオン相関関数を測定する。作用を定義する変数だけでなく、測定する量もゲージ不変に構成しなければならない。

A7 Wilson 作用による数値計算

Wilson は 1974 年、リンク変数と Wilson loop を用いることで、格子上で局所ゲージ対称性を保った非摂動的定式化を与えた [8]。強結合展開では大きな loop に対して

$$\langle W(C) \rangle \propto e^{-\sigma A(C)} \quad (225)$$

という面積則が現れる。 $A(C) = RT$ なら $\langle W(R, T) \rangle \sim e^{-\sigma RT}$ であり、式 $\langle W(R, T) \rangle \sim e^{-V(R)T}$ との比較から

$$V(R) \simeq \sigma R \quad (226)$$

という線形ポテンシャルが得られる。 σ は string tension である。

その後、Creutz は SU(2) 格子ゲージ理論を Monte Carlo で調べ、格子ゲージ理論が数値実験として実行できることを具体的に示した [15]。Wilson 作用は最も単純な離散化であり、現代の計算では離散化誤差を減らす改良作用も使われる。しかし、群に値を取るリンク変数を閉路に沿って掛けて作用を作り、Haar 測度で積分するという骨格は変わらない。

A8 グラスマン積分とフェルミオン行列式

格子上でフェルミオン作用は

$$S_f = \sum_{x,y} \bar{\psi}_x D_{xy}[U] \psi_y \quad (227)$$

という双線形になる。グラスマン積分を解析的に実行すると

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\bar{\psi}D[U]\psi} = \det D[U]. \quad (228)$$

2成分の例なら

$$\int d\bar{\psi}_2 d\psi_2 d\bar{\psi}_1 d\psi_1 \exp(-\bar{\psi}_i D_{ij}\psi_j) \quad (229)$$

$$= D_{11}D_{22} - D_{12}D_{21} = \det D \quad (230)$$

となる。グラスマン変数では各変数の2乗が0なので指数関数の展開は有限項で止まり、全変数を1度ずつ含む係数だけが積分後に残る。この係数が行列式である。従ってQCDの分配関数は

$$Z = \int \mathcal{D}U e^{-S_g[U]} \prod_{f=1}^{N_f} \det D_f[U]. \quad (231)$$

この行列式のゲージ場依存性まで配位生成に含めるクォークを、動的クォークと呼ぶ。摂動論の言葉では、閉じたクォークループによる真空偏極を含めることに対応する。行列式を定数とみなして省く近似はクエンチ近似と呼ばれる。グラスマン変数としてのフェルミオン自由度は積分後に消えたが、その効果は全リンクに依存する巨大疎行列の行列式として残る。

A9 擬フェルミオン表示

同じ質量を持つ2つの縮退したフレーバーを考える。両者のDirac演算子を $D_1 = D_2 = D$ と書けば、フェルミオン積分から現れる行列式の積は

$$\det D_1 \det D_2 = (\det D)^2 \quad (232)$$

となる。Wilson-Dirac 演算子などが γ_5 -Hermiticity $D^\dagger = \gamma_5 D \gamma_5$ を満たすとき、 $D^\dagger$ は D と相似であるため $\det D^\dagger = \det D$ であり、 $\det D$ は実数である。従って

$$(\det D)^2 = \det D^\dagger \det D = \det(D^\dagger D) \quad (233)$$

と変形できる。 $D^\dagger D$ は半正定値であり、零モードを避ける通常のシミュレーション条件では正定値として扱える。このとき

$$\det(D^\dagger D) \propto \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp\left[-\phi^\dagger (D^\dagger D)^{-1} \phi\right] \quad (234)$$

と書ける。この恒等式は複素ガウス積分

$$\int d\phi^\dagger d\phi e^{-\phi^\dagger A^{-1} \phi} \propto \det A \quad (235)$$

に $A = D^\dagger D$ を代入したものである。固有値分解 $A = V \text{diag}(\lambda_i) V^\dagger$ を行えば、各固有モードで

$$\int d\phi_i^* d\phi_i e^{-|\phi_i|^2/\lambda_i} \propto \lambda_i \quad (236)$$

となり、その積が $\prod_i \lambda_i = \det A$ を与えることが分かる。

ϕ は複素ボソン場だが、作用は非局所的である。力を計算するたびに

$$(D^\dagger D)\chi = \phi \quad (237)$$

のような巨大線形方程式を解く必要があり、動的フェルミオン計算の主要コストになる。擬フェルミオン場は、行列式を局所的な自由場へ置き換えたものではない。逆行列 $(D^\dagger D)^{-1}$ を通じて全格子点を結びつけるため、背景ゲージ場のもとでのフェルミオンの運動・伝播は、巨大な連立一次方程式として表現される。この連立一次方程式は反復型の線形ソルバーによって解かれ、作用、クォーク質量、格子間隔、アルゴリズムや実装に依存するものの、動的フェルミオンを含む格子計算では、しばしば全計算時間の50–90%程度を占める。

A10 重要度サンプリングとマルコフ連鎖 Monte Carlo 法

積分領域を一様に走査する代わりに、被積分関数への寄与が大きい $\pi(x) \propto e^{-S(x)}$ の領域を重点的に標本化する考え方を重要度サンプリングという。ここで単純 Monte Carlo 法とは、扱いやすい分布 $r(x)$ から全配位 x_i を独立に生成し、

$$\langle O \rangle = \frac{\mathbb{E}_r[O(x)w(x)]}{\mathbb{E}_r[w(x)]}, \quad w(x) = \frac{e^{-S(x)}}{r(x)} \quad (238)$$

として重み付き平均を計算する方法を指す。一様分布を使う場合もこの一例である。式自体は正しいが、連続場の積分領域は非有界なので一様分布をそのまま定義できず、有限範囲へ切っても高次元空間の大部分は作用が大きく、 e^{-S} への寄与がほとんどない。一般の r を使っても、 r の典型的な配位と目標分布 π の典型的な配位が重ならないければ、ごく少数の標本だけが巨大な w を持つ。このオーバーラップ問題により重みの分散が増え、名目上の標本数が多くても有効標本数は急速に小さくなる。

運動項がなければ重みは格子点ごとの積に因数分解するので、各点を独立に生成できる。しかし運動項から生じるホッピング項 $-\sum_{n,\mu} \varphi_n \varphi_{n+\hat{\mu}}$ があると、全配位の分布は因数分解しない。このとき各点を独立に提案して全配位を一度に置き換えると、作用差は一般に体積 V に比例して増え、受容率は模式的に $e^{-O(V)}$ まで低下しうる。

運動項があっても MCMC は利用できる。例えば 1 点 n だけを更新すれば、作用差 ΔS に寄与するのは φ_n の局所ポテンシャルと隣接点とのホッピング項だけであり、近傍の場を使って受容率を計算できる。この局所更新を反復すると、点ごとには独立でなくても、ホッピング項を含む全作用 $S_{\text{lat}}[\varphi]$ を定常分布とする相関した全配位を生成できる。後で導入する HMC 法は、全自由度を同時に動かしながら同じ目標分布を保つ、より大域的な更新法である。

このように格子場の高次元分布から独立標本を直接生成することは難しいため、 π を定常分布として保つ遷移確率を設計し、現在の配位から次の配位への更新を反復して相関した標本を作る。これがマルコフ連鎖 Monte Carlo 法であり、規格化定数 Z は確率の比で消える。具体的な Metropolis 法は小節 6.6 で導入する。有限状態の模型における遷移の反復と定常分布への収束は、本文の小節 6.5 で説明した。

A11 局所ランダムウォークの限界

高次元空間で各変数を少しずつランダムに動かすと、典型集合に沿って進むのが遅い。HMC 法は作用の勾配を使い、受容率を保ちながら遠くへ移動する。

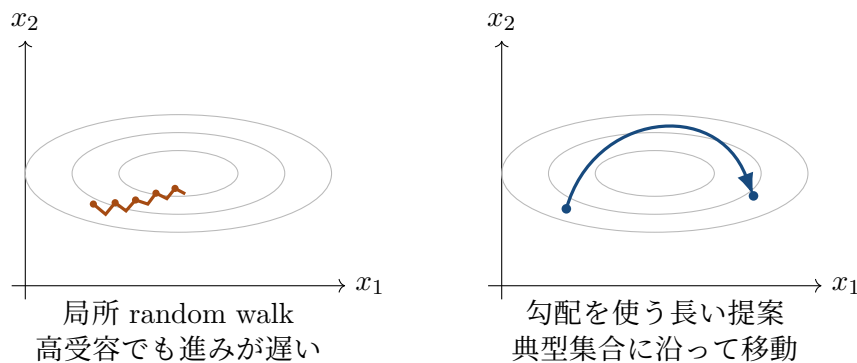


図 18: 高次元分布での局所提案と HMC 型提案の模式比較。性能は受容率だけでなく、配位空間をどれだけ移動し自己相関を減らすかで決まる。

A12 HMC 法が目標分布を保つ理由

式 (127) の拡張分布は、運動量を積分すると $e^{-S(\phi)}$ を周辺分布として与える。残る問題は、有限刻み幅の数値積分で作った提案がこの分布を厳密に保つかどうかである。

リープフロッグ積分が定める写像を $F : (\phi, p) \mapsto (\phi', p')$ とする。HMC で必要な性質は次の 2 つである。

可逆性 運動量反転 $R(\phi, p) = (\phi, -p)$ を用いて $F^{-1} = RFR$ と書ける。

体積保存 ヤコビアンが $|\det(\partial(\phi', p')/\partial(\phi, p))| = 1$ である。

リープフロッグ積分を構成する各部分更新はせん断変換であり、例えば

$$p' = p - \frac{\epsilon}{2} \nabla S(\phi), \quad \left| \det \frac{\partial(\phi, p')}{\partial(\phi, p)} \right| = 1, \quad (239)$$

$$\phi' = \phi + \epsilon M^{-1} p', \quad \left| \det \frac{\partial(\phi', p')}{\partial(\phi, p')} \right| = 1. \quad (240)$$

したがって、それらの合成であるリープフロッグ積分全体も体積を保存する。

可逆で体積保存な決定論的提案に対し、順方向と逆方向の密度比は

$$\frac{\rho(\phi', p')}{\rho(\phi, p)} = e^{-H_{\text{HMC}}(\phi', p') + H_{\text{HMC}}(\phi, p)} = e^{-\Delta H_{\text{HMC}}} \quad (241)$$

である。従って式 (298) は、拡張位相空間に対する Metropolis–Hastings 受容率そのものである。つまり HMC の正しさは

$$\text{運動量の正しい再生成} + \text{可逆性} + \text{体積保存} + \text{Metropolis 補正} \quad (242)$$

の組合せで成り立つ。単に ΔH_{HMC} が小さいだけでは十分ではない。

シンプレクティック積分とエネルギー誤差の関係は付録 A34 にまとめる。

A13 ゲージ場に対する HMC 法

SU(3) リンクは線形空間の変数ではない。運動量 $P_\mu(n)$ を Lie 代数 $\mathfrak{su}(3)$ に取り、

$$H_{\text{HMC}}[U, P] = S_g[U] + S_{\text{pf}}[U, \phi] + \frac{1}{2} \sum_{n, \mu} \text{Tr } P_\mu(n)^2 \quad (243)$$

とする。リンク更新は

$$U_\mu(n) \longrightarrow e^{i\epsilon P_\mu(n)} U_\mu(n) \quad (244)$$

の形になり、常に SU(3) 群上に留まる。ゲージ作用の力は周囲のステーブルから、擬フェルミオン力は Dirac 演算子の逆作用から得る。

A14 コードとの対応

実装では次の層を分けると理解しやすい。

1. 幾何: 格子サイズ、周期境界条件、前後の隣接点。
2. 場: サイト場、リンク場、Lie 代数に値を取る運動量。
3. 作用: プラケット和、フェルミオン行列、擬フェルミオン作用。
4. 微分: ゲージ力、フェルミオン力。
5. 積分器: リープフロッグ積分や Omelyan 積分器。
6. マルコフ更新: 運動量生成、軌道、Metropolis 判定、測定。

Julia の LatticeQCD.jl では、利用者はパラメータを設定して計算を走らせるが、内部ではこの順序が実行される。

```
using LatticeQCD
run_wizard()           # parameter file generator
run_LQCD("my_parameters.toml") # configuration generation
```

コードを読むときは、各処理がどの数式に対応するかを追うとよい。

A15 ニューラルネットを用いた確率分布の表現

ニューラルネットをパラメータ θ を持つ関数 f_θ と見る。教師あり学習では入力 x と正解 y に対して

$$\theta_* = \arg \min_{\theta} \frac{1}{N} \sum_i \ell(f_\theta(x_i), y_i) \quad (245)$$

を解く。一方、生成モデルでは関数値より分布 $q_\theta(x)$ を学ぶ。

格子場理論で必要なのは、ニューラルネット一般の分類能力ではなく、高次元の場の間に計算可能な変換を作る能力である。後に述べるアフィンカップリング層では、ニューラルネット自身が配位を直接出力するのではなく、可逆変換のスケール s_θ と平行移動 t_θ を出力する。この構成により、ネットワーク内部が複雑でも、配位全体の変換は可逆であり、その確率密度を評価できる。

ここでの目的は、目標分布

$$\pi(x) = \frac{e^{-S(x)}}{Z} \quad (246)$$

に近く、かつサンプルと密度評価が容易な $q_\theta(x)$ を作ることである。そのため、モデルには少なくとも次の3つの操作が必要になる。

1. 基底分布から新しい配位を生成すること。
2. 与えられた配位のモデル密度 $q_\theta(x)$ を評価すること。
3. 損失関数を θ で微分し、勾配法で学習できること。

画像生成では密度評価を省く手法も有力だが、Metropolis 補正を行うには $q_\theta(x)$ の値または提案確率の比が必要である。

A16 ヤコビアン計算の課題と設計上の工夫

格子体積を V 、格子点ごとの自由度を c とすれば、場の次元は $N = cV$ である。一般の写像 $f: \mathbb{R}^N \rightarrow \mathbb{R}^N$ のヤコビアンは $N \times N$ の密行列であり、行列を保持するだけでも $O(V^2)$ 、通常の LU 分解で行列式を求めれば $O(V^3)$ の計算量が必要になる。これは大体積の格子計算では受け入れられない。しかも $\log |\det J_f|$ を捨てると変数変換後の確率密度を評価できず、MH 補正も正しく構成できない。

そこで正規化フローでは、表現力を保ちながら行列式だけは厳密かつ安価に計算できるよう、写像の構造そのものを制限する。小さな写像 f_k の合成 $f = f_K \circ \dots \circ f_1$ を用いれば

$$\log |\det J_f| = \sum_{k=1}^K \log |\det J_{f_k}| \quad (247)$$

と分解できる。さらに各 J_{f_k} を三角行列または block 三角行列にする、局所的で疎な変換を使う、連続 flow で対数行列式を trace 積分へ置き換える、といった工夫が用いられる。要点は、行列式の正確な値を安く得られる構造を最初から設計することにある。代表的なアフィンカップリング層の具体式は付録で説明する。

A17 正規化フローの発展と格子場理論への応用

正規化フローは、可逆変換によって単純な分布を柔軟な分布へ変える方法として変分推論の文脈で体系化された [25]。アフィンカップリング層を用いて高次元でも逆変換とヤコビアンを安価にした代表例が Real NVP である [26]。実装上の詳細は付録に回し、ここではこれらが式 (154) をニューラルネットで近似するための計算可能な写像を与えることだけを押さえる。

Albergo らは ϕ^4 理論でフローを独立提案として用い、MH 補正によって厳密なマルコフ連鎖を作る形を示した [33]。その後、ゲージ対称性を層ごとに保つフローが $U(1)$ および $SU(N)$ へ拡張され [37, 38]、擬フェルミオンを含む理論への構成も研究されている [39]。この発展から、表現力だけでなく対称性、密度評価、分布の台、体積・相関長に対するスケーリングが中心課題であることが分かる。

A18 1 変数変換の幾何学的理解

$z \sim \mathcal{N}(0, 1)$ とし、単調な $x = f(z)$ を考える。確率保存

$$q(x)dx = r(z)dz \quad (248)$$

から

$$q(x) = r(z) \left| \frac{dz}{dx} \right| \quad (249)$$

となる。 f が空間を引き伸ばす場所では密度は薄くなり、押し縮める場所では濃くなる。多変数では伸縮率がヤコビアンの行列式になる。

A19 ゲージ共変フロー

格子ゲージ場では、通常の画像用ネットワークをリンク行列へそのまま適用するとゲージ対称性を壊す。設計の選択肢は

1. ゲージ固定後の変数を学習する、
2. ゲージ不変な閉路だけで変換を作る、
3. 層そのものをゲージ共変にする

ことである。ゲージ変換した配位を U^Ω と書くと、ゲージ共変な写像は

$$f_\theta(U^\Omega) = f_\theta(U)^\Omega \quad (250)$$

を満たす。従って、同じゲージ軌道に属する入力、変換後も同じゲージ軌道へ写される。この性質を層ごとに満たせば、学習中や生成後にゲージ対称性を近似的に回復させる必要がない。

有限 a で厳密な対称性を保つことは、単なるデータ拡張より強い帰納バイアスになる。 $SU(N)$ リンクを直接扱うゲージ共変フローの具体的構成も提案されている [37, 38]。

ゲージ共変な学習写像の具体例も見ておこう。永井・富谷のゲージ共変ニューラルネットワークは、stout smearing を学習可能な残差写像として書き直す [34]。模式的には

$$U_\mu^{(\ell+1)}(n) = \exp \left[i \sum_f \rho_{\ell f} Q_\mu^{(\ell, f)}(n) \right] U_\mu^{(\ell)}(n), \quad (251)$$

であり、 $Q_\mu^{(\ell, f)}(n) \in \mathfrak{su}(N)$ を staple やより大きな閉路から作り、 $\rho_{\ell f}$ を学習する。 Q_μ が格子点 n で随伴変換するため、出力リンクも入力リンクと同じゲージ変換則を満たす。この構成では既存の smearing とその力の計算法を、学習可能なネットワークと backpropagation として解釈できる。

Ipp らの lattice gauge equivariant convolutional neural network (L-CNN) は、近傍の特徴量をリンクに沿って同じ格子点まで平行移動してから畳み込み、同じ点に集めた共変な行列値特徴量を双線

形層で掛け合わせる [35]。層を重ねるとさまざまな形の Wilson loop を構成でき、ゲージ共変な関数を系統的に表現できる。

永井・大野・富谷の CASK (Covariant Attention with Stout Kernel) は、この stout 型共変写像へ Transformer の attention を導入した [36]。link と extended staple が同じ両端を持つように組み、両者の Frobenius 内積からゲージ不変な attention 係数を作る。その係数で共変な link 情報を混合するため、出力はゲージ共変性を保ちつつ、離れた link 間の相関も取り込める。

これらはゲージ共変な写像や作用モデルを作るためのアーキテクチャである。ただし、すべてがそのまま正規化フローになるわけではない。密度を評価するフローとして使うには、ゲージ共変性に加えて、写像の可逆性と Jacobian を計算できる構造が必要である。

A20 HMC 法との組み合わせ

Lüscher の提案 HMC 法とフローを組み合わせる原型は、Lüscher による trivializing map の構成にある [21]。物理的な場 x を可逆写像 $x = f(z)$ で表すと、潜在変数 z の作用は

$$S_{\text{eff}}(z) = S(f(z)) - \log |\det J_f(z)| \quad (252)$$

である。従って、元の作用 $S(x)$ で HMC 法を行う代わりに

$$H_{\text{flow}}(z, p) = \frac{1}{2} p^T p + S_{\text{eff}}(z) \quad (253)$$

を用いて z 空間で HMC 法を実行し、得られた z' を $x' = f(z')$ へ戻せばよい。Lüscher は、場の空間の flow を積分してこの写像を構成し、Wilson flow から得られる近似 trivializing map を HMC 法と組み合わせることを提案した。

写像が目標分布を完全に trivialize すれば S_{eff} は簡単な基底作用になり、異なる自由度間の強い相関は消える。実際には近似写像で十分であり、 S_{eff} の曲率や長距離相関を弱められれば、潜在空間の HMC 軌道は元の変数より移動しやすくなる。リープフロッグ積分の最後に

$$A = \min(1, e^{-\Delta H_{\text{flow}}}) \quad (254)$$

で補正すれば、写像が不完全でも定常分布は変わらない。ただし、作用の力だけでなく $\partial_z \log |\det J_f|$ も毎ステップ計算するため、写像による自己相関の短縮と追加計算量を壁時計時間で比較する必要がある。

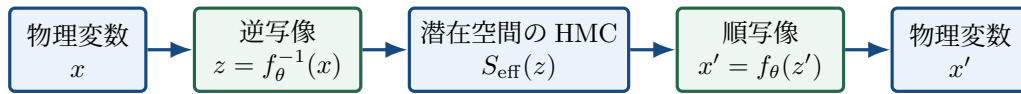


図 19: field-transformed HMC または flowed HMC の概念図。HMC 法を、写像で単純化した潜在空間で実行する。

機械学習による写像 Foreman、Izubuchi、Jin らは、Lüscher の field transformation を学習可能な normalizing flow で実装した [23]。この方法では f_θ を式 (252) へそのまま入れ、潜在空間の分子動力学で $S(f_\theta(z)) - \log |\det J_{f_\theta}(z)|$ の力を使う。2次元 U(1) ゲージ理論の proof of concept では、通常の HMC 法ではほぼ固定されていたトポロジカルチャージが複数のセクターを移動し、畳み込み型のゲージ共変層の重みを大きな体積へ移して少ない再学習で利用できることも示された。この構成ではフローが HMC 法を置き換えるのではなく、HMC 法が動く座標系を学習する。

Schwinger–Dyson 方程式による写像 写像の設計はニューラルネットに限らない。Boyle、Izubuchi、Jin らは、flow kernel を有限個の Wilson loop の基底で展開し、その係数を Schwinger–Dyson 方程式と格子上の期待値から決める近似 trivializing map を構成した [24]。解析的な flow-time 展開だけに頼らず、選んだ演算子空間で有効作用を制御できる点が特徴である。変換後の正確な作用で HMC 法を

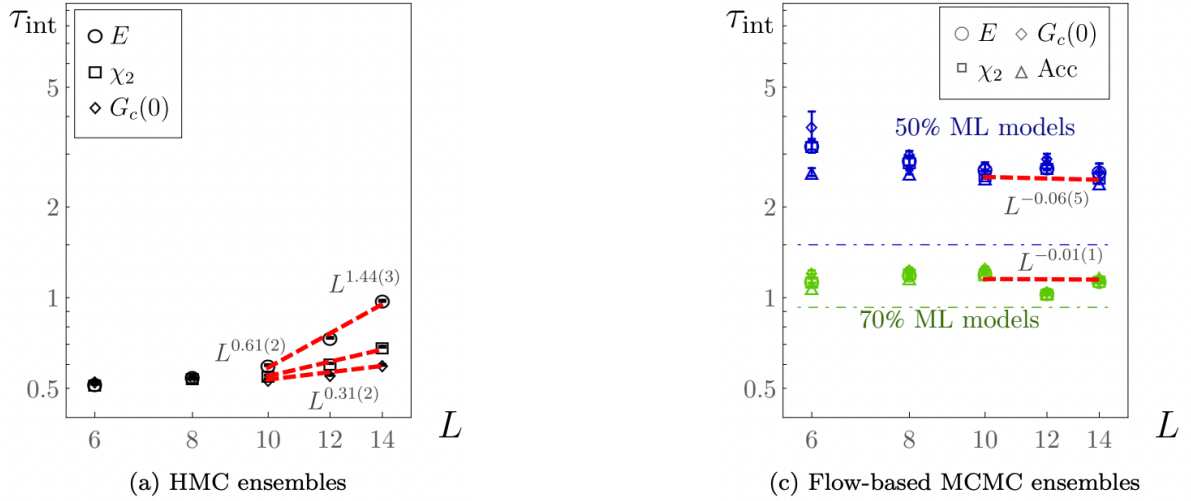


図 20: 2次元スカラー場における自己相関時間の体積依存性の例。左はHMC法、右はフローを用いたMCMC。これは小規模ベンチマークであり、4次元QCDで同じ改善が保証されるわけではない。

行くと、長距離観測量の decorrelation が速くなる場合があった一方、拡張 Wilson loop、ヤコビアン、Hessian の計算費用が利得を上回った。従って、良い写像とは S_{eff} を平坦にする写像だけでなく、その力とヤコビアンを安価に評価できる写像でなければならない。

このほかにも、フローによる独立提案とHMC更新を交互に行う方法や、学習した積分器をMH補正と組み合わせる方法がある。共通する考え方は、学習器または近似写像には移動効率を改善させ、最終的な受容判定には元の理論の正確な作用を使うことである。

A21 サンプラーの評価基準

生成画像の見た目に相当する評価だけでは不十分である。格子場理論のサンプラーは次で評価する。

1. 厳密性: MH 補正後に正しい定常分布を持つか。
2. 被覆性: 重要なモードやトポロジカルセクターを覆うか。
3. 計算効率: 実行時間あたりに何個の有効な独立標本が得られるか。
4. スケーリング: 体積増大、軽い質量、連続極限に対して性能がどう変化するか。
5. 学習コストの償却: 学習に要した計算量を何配位の生成で回収できるか。

1つの観測量 O について、実行時間 t_{run} あたりの情報量は

$$\frac{N_{\text{eff}}(O)}{t_{\text{run}}} \simeq \frac{N}{2\tau_{\text{int}}(O)t_{\text{run}}} \quad (255)$$

で比較できる。観測量ごとに自己相関時間が異なるため、平均受容率だけで順位を決めてはならない。小格子で高受容率を得ることと、QCDの連続極限でHMC法を上回ることは大きな隔たりがある。

A22 絶対零度と松原形式

ここまで現れた β は、熱場の理論では自然単位系 $k_B = 1$ を用いて $\beta = 1/T$ と解釈できる。トレースを取るため虚時間の両端 $\tau = 0, \beta$ が同一視され、Euclid 時間方向は長さ β の円になる。ボース場は周期境界条件、フェルミオン場は反周期境界条件

$$\phi(\tau + \beta, \mathbf{x}) = \phi(\tau, \mathbf{x}), \quad \psi(\tau + \beta, \mathbf{x}) = -\psi(\tau, \mathbf{x}) \quad (256)$$

を満たす。このため虚時間方向の Fourier 周波数は連続ではなく、

$$\omega_n^B = 2\pi nT, \quad \omega_n^F = (2n + 1)\pi T, \quad n \in \mathbb{Z} \quad (257)$$

となる。これが松原形式である。 $T \rightarrow 0$ では $\beta \rightarrow \infty$ となり、周波数の間隔が詰まって $T \sum_n$ は連続な周波数積分 $\int d\omega/(2\pi)$ へ移る。

絶対零度で真空情報が得られる理由は、Euclid 時間発展が低エネルギー状態への射影として働くためである。エネルギーの原点を固定して E_0 を基底状態エネルギーとすると、

$$Z(\beta) = \sum_n e^{-\beta E_n} = e^{-\beta E_0} \left[1 + \sum_{n>0} e^{-\beta(E_n - E_0)} \right] \quad (258)$$

であり、 $\beta \rightarrow \infty$ では励起状態の寄与が指数的に消える。真空が一意なら

$$\langle O \rangle_\beta = \frac{\text{Tr}(e^{-\beta H} O)}{Z(\beta)} \xrightarrow{\beta \rightarrow \infty} \langle 0|O|0 \rangle \quad (259)$$

となる。同様に $0 < \tau < \beta$ で

$$\begin{aligned} G_\beta(\tau) &= \frac{1}{Z(\beta)} \text{Tr} \left(e^{-(\beta-\tau)H} O e^{-\tau H} O \right) \\ &\xrightarrow{\beta \rightarrow \infty} \langle 0|O e^{-\tau(H-E_0)} O|0 \rangle \end{aligned} \quad (260)$$

となる。従って、十分長い Euclid 時間方向を持つ格子で測る相関関数から、真空期待値、励起エネルギー、行列要素を取り出せる。真空が縮退している場合には、境界条件や微小な外場などで対象とするセクターを選ぶ必要がある。

A23 スペクトル表示と実時間再構成

例えば零温度ではスペクトル密度 ρ を用いて

$$G_E(\tau, \mathbf{p}) = \int_0^\infty d\omega \rho(\omega, \mathbf{p}) e^{-\omega\tau}, \quad G_M(t, \mathbf{p}) = \int_0^\infty d\omega \rho(\omega, \mathbf{p}) e^{-i\omega t} \quad (261)$$

と核が異なる。有限個かつ統計誤差を含む Euclid 時間データから ρ を逆算するのは悪条件の逆 Laplace 問題であるため、時間的相関や光円錐上の相関は通常の Euclid 格子計算から直接求めることが難しい。再構成定理は原理的な対応を保証するが、有限データからの数値的再構成を自動的に容易にするものではない。

A24 自由場と多変量ガウス分布

$\lambda_0 = 0$ なら作用は 2 次形式

$$S = \frac{1}{2} \sum_{n,m} \varphi_n K_{nm} \varphi_m = \frac{1}{2} \varphi^T K \varphi \quad (262)$$

である。したがって

$$p(\varphi) \propto \exp\left(-\frac{1}{2}\varphi^T K \varphi\right) \quad (263)$$

は共分散 K^{-1} の多変量ガウス分布である。格子上の隣接項が、異なる格子点の場を相関させる。相互作用 $\lambda_0 \varphi^4$ を入れると非ガウス分布になる。後半で扱う正規化フローは、簡単なガウス分布をこの複雑な分布へ変形する生成モデルである。

A25 自由場の運動量表示と格子伝播関数

周期境界条件のもとで格子 Fourier 変換を

$$\varphi_n = \frac{1}{\sqrt{V}} \sum_p e^{ip \cdot n} \tilde{\varphi}(p), \quad p_\mu = \frac{2\pi k_\mu}{N_\mu}, \quad k_\mu = 0, 1, \dots, N_\mu - 1 \quad (264)$$

と定義する。隣接差分は

$$\varphi_{n+\hat{\mu}} - \varphi_n = \frac{1}{\sqrt{V}} \sum_p e^{ip \cdot n} (e^{ip_\mu} - 1) \tilde{\varphi}(p) \quad (265)$$

となる。したがって

$$|e^{ip_\mu} - 1|^2 = 2 - 2 \cos p_\mu = 4 \sin^2 \frac{p_\mu}{2}. \quad (266)$$

式 (31) の自由作用は

$$S_0 = \frac{1}{2} \sum_p \tilde{\varphi}(-p) [\hat{p}^2 + (am_0)^2] \tilde{\varphi}(p), \quad (267)$$

$$\hat{p}^2 \equiv 4 \sum_\mu \sin^2 \frac{p_\mu}{2} \quad (268)$$

となる。ガウス積分の逆行列が 2 点関数なので

$$\langle \tilde{\varphi}(p) \tilde{\varphi}(-p) \rangle = \frac{1}{\hat{p}^2 + (am_0)^2}. \quad (269)$$

小さい p_μ では

$$\hat{p}^2 = \sum_\mu p_\mu^2 + O(p^4) \quad (270)$$

だから、連続理論の伝播関数 $(p^2 + m_0^2)^{-1}$ へ戻る。一方、ブリルアンゾーンの端では $\hat{p}^2$ は p^2 と大きく異なり、これが有限格子間隔の離散化効果である。

位置空間の 2 点関数は

$$G(n) = \frac{1}{V} \sum_p \frac{e^{ip \cdot n}}{\hat{p}^2 + (am_0)^2} \quad (271)$$

である。長距離では $G(n) \sim e^{-|n|/\xi}$ と減衰し、自由理論では $\xi \simeq 1/(am_0)$ である。 $a \rightarrow 0$ で物理質量 m_0 を固定すれば $am_0 \rightarrow 0$ 、すなわち格子単位の相関長 $\xi \rightarrow \infty$ となる。この式は、量子連続極限が臨界点への接近として見える理由を具体化している。

格子上の摂動論 格子場理論は、連続時空の方程式を Monte Carlo 法で数値計算するためだけの手法ではなく、紫外カットオフを備えた場の量子論そのものである。自由格子伝播関数と局所相互作用から Feynman 則を定め、ブリルアンゾーン内の運動量として摂動論を組み立てることもできるほか、強結合展開、ホッピング・パラメータ展開、転送行列、くりこみ群などの解析的手法も使える。格子化は理論を曖昧さのない有限自由度の問題として定義し、解析計算と数値計算の双方を可能にする。格子上の摂動展開と 1 ループ補正の具体例は付録 A27 に示す。

A26 Gell–Mann–Low 方程式

式 (39) は、くりこみ条件に選んだ期待値に対する全微分の式であった。これを n 点相関関数に特定し、場の規格化の変化も含めたものが Gell–Mann–Low 方程式である [11]。ここでは、くりこみ点 μ を動かす表示ではなく、紫外カットオフ Λ を変える表示を用いる。

紫外カットオフを入れて計算した n 点関数を $G_\Lambda^{(n)}(p_1, \dots, p_n; m, \lambda)$ と書く。外線の運動量 p_i を物理単位で固定する。裸の場 ϕ と裸の相関関数を場のくりこみ定数 $Z_\phi(\Lambda)$ で規格化し、

$$\phi_R = Z_\phi^{-1/2}(\Lambda)\phi, \quad G_R^{(n)}(p_1, \dots, p_n) = Z_\phi^{-n/2}(\Lambda)G_\Lambda^{(n)}(p_1, \dots, p_n; m(\Lambda), \lambda(\Lambda)) \quad (272)$$

と定義する。等号の右辺は Λ に明示的に依存するが、 $m(\Lambda)$ 、 $\lambda(\Lambda)$ 、 $Z_\phi(\Lambda)$ を同時に調整し、左辺が同じ物理的相関関数を表すようにする。そこで、式 (39) のパラメータ変化率と、場の異常次元を

$$\beta_\lambda^{(\Lambda)}(m, \lambda) = \Lambda \frac{d\lambda}{d\Lambda}, \quad \beta_m^{(\Lambda)}(m, \lambda) = \Lambda \frac{dm}{d\Lambda}, \quad \gamma_\phi^{(\Lambda)}(m, \lambda) = -\frac{1}{2}\Lambda \frac{d \ln Z_\phi}{d\Lambda} \quad (273)$$

と定義する。最後の符号は、以下の式で異常次元項を正符号で書くための規約である。外線運動量 p_i を固定したまま、式 (272) を Λ で全微分する。左辺は物理的相関関数として不変に保つので、

$$\begin{aligned} 0 &= \Lambda \frac{d}{d\Lambda} G_R^{(n)}(p_1, \dots, p_n) \Big|_{p_i \text{ fixed}} \\ &= \Lambda \frac{d}{d\Lambda} \left[Z_\phi^{-n/2}(\Lambda) G_\Lambda^{(n)} \right] \\ &= Z_\phi^{-n/2}(\Lambda) \left[\Lambda \frac{\partial}{\partial \Lambda} + \beta_\lambda^{(\Lambda)}(m, \lambda) \frac{\partial}{\partial \lambda} + \beta_m^{(\Lambda)}(m, \lambda) \frac{\partial}{\partial m} + n\gamma_\phi^{(\Lambda)}(m, \lambda) \right] G_\Lambda^{(n)}. \end{aligned} \quad (274)$$

最後の等号では、式 (273) を用いて、明示的な Λ 依存性と m, λ, Z_ϕ を通じた暗黙的な依存性を分けて書いた。共通因子 $Z_\phi^{-n/2}$ を除けば、カットオフ表示の Gell–Mann–Low 方程式は

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta_\lambda^{(\Lambda)} \frac{\partial}{\partial \lambda} + \beta_m^{(\Lambda)} \frac{\partial}{\partial m} + n\gamma_\phi^{(\Lambda)} \right] G_\Lambda^{(n)} = 0 \quad (275)$$

となる。 Λ だけを動かせば、積分に含める短距離自由度が変わるため $G_\Lambda^{(n)}$ も変わる。しかし $m(\Lambda)$ 、 $\lambda(\Lambda)$ と場の規格化を同時に動かせば、同じ長距離の物理を表し続ける。

たとえば 4 次元の $\lambda\phi^4/4!$ 理論では、結合定数の流れは 1 ループで

$$\beta_\lambda^{(\Lambda)}(\lambda) = \frac{3\lambda^2}{16\pi^2} + O(\lambda^3), \quad \gamma_\phi^{(\Lambda)}(\lambda) = O(\lambda^2). \quad (276)$$

となる。一方、紫外カットオフを明示した正則化では、質量 2 乗には $\lambda\Lambda^2$ に比例する加法的な補正が生じる。そのため質量の流れは一般に $\gamma_m m$ という乗法的な項だけでは表せず、臨界質量からのずれを Λ ごとに調整する必要がある。この調整が、格子上で連続極限へ近づくときの質量パラメータの微調整に対応する。

A27 格子上の摂動論

相互作用を加えた理論も、連続時空と同様に自由理論のまわりで摂動展開できる。 $S = S_0 + S_{\text{int}}$ 、

$$S_{\text{int}} = \frac{\lambda_0}{4!} \sum_n \varphi_n^4 \quad (277)$$

と分ければ、観測量の期待値は

$$\begin{aligned} \langle O \rangle &= \frac{\langle O e^{-S_{\text{int}}} \rangle_0}{\langle e^{-S_{\text{int}}} \rangle_0} \\ &= \langle O \rangle_0 - \langle O S_{\text{int}} \rangle_{0,c} + \frac{1}{2} \langle O S_{\text{int}}^2 \rangle_{0,c} + \cdots \end{aligned} \quad (278)$$

と展開される。添字 0 は自由格子作用によるガウス平均、 c は連結部分を表す。Wick の定理を使えば、式 (269) が Feynman 図の内部線となり、 φ^4 項が 4 点頂点になる。連続理論の運動量積分に代わって、各ループ運動量はブリルアンゾーン内の格子運動量について和を取る。例えば 1 ループのタドポールによる質量補正は

$$\delta(am)^2_{1\text{-loop}} = \frac{\lambda_0}{2V} \sum_k \frac{1}{\widehat{k}^2 + (am_0)^2} \quad (279)$$

の形になる。有限の a では運動量領域が有限なので紫外発散は現れないが、 $a \rightarrow 0$ では高運動量モードが再び増え、裸の質量や結合を調整するくりこみが必要になる。

このように、格子場理論は連続時空の方程式を数値計算する手法の名前ではなく、離散時空上に定義された場の量子論である。その上では Monte Carlo 法だけでなく、摂動論、強結合展開、ホッピング・パラメータ展開、転送行列、くりこみ群などの解析的・数値的な道具を使える。格子化の重要な点は、場の理論を曖昧さのない有限自由度の問題として定義し、目的に応じて複数の方法で調べられることである。

A28 ダブラーが連続極限で消えない理由

余分な零点は $p_\mu = \pi/a$ 付近にあるため、 $a \rightarrow 0$ では無限大の運動量へ移動し、低エネルギーの物理から消えるようにも見える。しかし、零点の絶対的な位置が発散することと、その零点の近傍に有限エネルギーの励起が残ることは別である。ブリルアンゾーンの各角を $\eta_\mu \in \{0, 1\}$ で指定し、その近傍の運動量を

$$p_\mu = \frac{\pi \eta_\mu}{a} + k_\mu \quad (280)$$

と書く。ここで残留運動量 k_μ は、 $a \rightarrow 0$ でも固定しておく物理的な運動量である。このとき

$$\begin{aligned} \frac{1}{a} \sin(ap_\mu) &= \frac{1}{a} \sin(\pi \eta_\mu + ak_\mu) \\ &= (-1)^{\eta_\mu} k_\mu + O(a^2 k_\mu^3) \end{aligned} \quad (281)$$

なので、各角の近傍でナイーブ Dirac 演算子は

$$D_{\text{naive}} \left(\frac{\pi \boldsymbol{\eta}}{a} + \mathbf{k} \right) = m + i \sum_\mu \gamma_\mu^{(\boldsymbol{\eta})} k_\mu + O(a^2 k^3), \quad \gamma_\mu^{(\boldsymbol{\eta})} = (-1)^{\eta_\mu} \gamma_\mu \quad (282)$$

となる。符号を変えた $\gamma_\mu^{(\boldsymbol{\eta})}$ も同じ Clifford 代数を満たす。従って伝播関数は

$$S_{\boldsymbol{\eta}}(\mathbf{k}) = D_{\text{naive}}^{-1} \left(\frac{\pi \boldsymbol{\eta}}{a} + \mathbf{k} \right) \xrightarrow{a \rightarrow 0} \left[m + i \sum_\mu \gamma_\mu^{(\boldsymbol{\eta})} k_\mu \right]^{-1} \quad (283)$$

となり、すべての角から質量 m の連続 Dirac フェルミオンが現れる。すなわち、逆伝播関数の零点は伝播関数の有限質量の極として残る。零点の中心 $\pi\eta/a$ は無限遠へ移るが、そのまわりの有限な k を持つ領域は消えない。実際、ブリルアンゾーンの運動量積分は各角の近傍に分割でき、それぞれが連続極限の外線状態やループ積分へ有限に寄与する。

したがって、余分な極が現れても、それに対応する質量が $1/a$ のように発散して伝播関数が低エネルギーで消えるなら問題はない。しかし、ナイブフェルミオンのダブラーは質量 m のままであり、単に極の中心がカットオフ付近にあるだけなので分離しない。Wilson 項は、ダブラーへ $O(1/a)$ の質量を与えることで、この条件を実現する。

A29 point splitting 法とゲージ不変な複合演算子

同じ点にある場から作る局所複合演算子 $\bar{\psi}(x)\Gamma\psi(x)$ は、内部指標をゲージ重項に縮約すればゲージ不変にできる。一方、紫外特異性を調べるために場をわずかに異なる 2 点へ分離し、

$$\bar{\psi}\left(x+\frac{\epsilon}{2}\right)\Gamma\psi\left(x-\frac{\epsilon}{2}\right) \quad (284)$$

とすると、両端の場は異なるゲージ変換行列で変換するため、この積だけではゲージ不変でない。これは point splitting 法で複合演算子を定義するとき生じる基本的な問題である。

そこで、式 (55) の平行移動子を 2 点の間に挿入する。式 (56) を用いると、例えば点分離した電流

$$j_{\epsilon}^{\mu}(x)=\bar{\psi}\left(x+\frac{\epsilon}{2}\right)\gamma^{\mu}W_C\left(x+\frac{\epsilon}{2},x-\frac{\epsilon}{2}\right)\psi\left(x-\frac{\epsilon}{2}\right) \quad (285)$$

はゲージ不変になる。演算子の短距離発散を必要な反項で差し引いた後に $\epsilon \rightarrow 0$ を取ることで、局所複合演算子を定義する。

より一般に、異なる点 $x_1, \dots, x_n$ に置かれた複数の場から複合演算子を作る場合には、各場を Wilson line で共通の基準点へ平行移動し、その点で内部指標をゲージ重項へ縮約する。有限の点分離では演算子は選んだ経路に依存し、局所極限では同じ量子数を持つ演算子との混合も考慮する必要がある。従って、Wilson line の挿入はゲージ不変性を保証するが、複合演算子のくりこみそのものを不要にするわけではない。

$U(1)$ 理論ではゲージ場が可換なので経路順序積は不要であり、式 (55) は Schwinger phase と呼ばれる位相因子になる。Schwinger はこの位相を用いて、ゲージ共変な量だけから真空偏極を扱う方法を明確にした [5]。非可換理論では経路順序積が必要になり、その一般化が Wilson line である。

A30 マルコフ連鎖の定常性とエルゴード性

マルコフ連鎖では現在の状態 x から遷移確率 $T(x \rightarrow x')$ で次の状態 x' を作る。目標分布を定常分布にする十分条件が詳細釣り合い

$$\pi(x)T(x \rightarrow x') = \pi(x')T(x' \rightarrow x) \quad (286)$$

である。両辺を x で積分すれば

$$\int dx \pi(x)T(x \rightarrow x') = \pi(x') \quad (287)$$

となる。詳細釣り合いは十分条件であって必要条件ではない。より一般には、確率流の総和が各点で釣り合う大域的釣り合いが成り立てば π は定常分布になる。詳細釣り合いは順方向と逆方向を 1 組ずつ等しくするため、設計と証明が容易であり、Metropolis–Hastings 法や HMC 法で広く用いられる。

定常分布が正しいことだけでは、任意の初期状態からそこへ到達できるとは限らない。状態空間のどこへも有限ステップで到達できる既約性と、特定の周期だけを往復しない非周期性が必要である。これらをまとめてエルゴード性と呼ぶことが多い。実際の計算では、初期配位の影響が残る熱化区間を測定から除き、異なる初期条件から得た連鎖が同じ分布へ近づくかを確認する。厳密な定常分布を持つアルゴリズムでも、熱化不足や有限標本数による誤差まで自動的に消えるわけではない。

A31 Metropolis–Hastings 受容率と詳細釣り合い

前節では、目標分布をマルコフ連鎖の定常分布にする条件として詳細釣り合いを導入し、さらに、その定常分布へ実際に到達するためにはエルゴード性が必要であることを述べた。ここでは、その条件を具体的に満たす更新規則として Metropolis–Hastings 法を考える。考え方は、まず proposal 分布 $q(x'|x)$ から候補 x' を作り、その候補を受容確率 $A(x \rightarrow x')$ で採用し、棄却された場合には元の状態 x に留まる、という二段階に分けることである。

この分け方をすると、候補をどう作るか、目標分布をどう保つかを別々に設計できる。proposal は局所的なランダムウォークでも、HMC の軌道でも、学習モデルからの提案でもよい。重要なのは、受容率に順方向と逆方向の proposal 確率の比を入れることで、前節の詳細釣り合いが成り立つようにする点である。以下でその確認を行う。

$x \neq x'$ に対する実際の遷移確率は

$$T(x \rightarrow x') = q(x'|x)A(x \rightarrow x') \quad (288)$$

である。式 (112) を代入すると

$$\pi(x)T(x \rightarrow x') = \pi(x)q(x'|x) \min \left[1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)} \right] \quad (289)$$

$$= \min [\pi(x)q(x'|x), \pi(x')q(x|x')] . \quad (290)$$

右辺は x と x' を入れ替えても変わらない。したがって

$$\pi(x)T(x \rightarrow x') = \pi(x')T(x' \rightarrow x) \quad (291)$$

が成り立つ。棄却確率

$$T(x \rightarrow x) = 1 - \int_{x' \neq x} dx' T(x \rightarrow x') \quad (292)$$

を加えれば確率は 1 に規格化される。

この証明から次の 2 点が分かる。第 1 に、 π の規格化定数は比で消える。第 2 に、受容率は経験的な罰則ではなく、順方向と逆方向の確率流を一致させるための係数である。

受容率が高いだけでは十分ではない。非常に小さな移動ならほぼ必ず受容されるが、配位空間を進まず自己相関が長くなる。

A32 学習モデルと厳密性の役割分担

ここで扱う学習モデルの基本的な役割は、目標分布を近似して物理結果を直接出すことではなく、現在の配位 x から次の候補 x' を作る proposal を改善することである。学習した提案核を $q_\theta(x'|x)$ と書けば、候補 x' を生成した後に

$$A_\theta(x \rightarrow x') = \min \left[1, \frac{\pi(x')q_\theta(x|x')}{\pi(x)q_\theta(x'|x)} \right] = \min \left[1, e^{-S(x')+S(x)} \frac{q_\theta(x|x')}{q_\theta(x'|x)} \right] \quad (293)$$

で受容・棄却する。これは第 6 節の Metropolis–Hastings 法を、学習した proposal へそのまま適用したものである。目標分布 $\pi(x) = e^{-S(x)}/Z$ の未知な規格化定数 Z は比の中で消える。順方向と逆方向の提案確率を正しく評価でき、必要な可逆性や分布の台が確保されていれば、式 (293) は π に関する詳細釣り合いを満たす。

特に、現在の配位によらず $x' \sim q_\theta(x')$ と生成する独立提案では

$$A_\theta(x \rightarrow x') = \min \left[1, \frac{e^{-S(x')}q_\theta(x)}{e^{-S(x)}q_\theta(x')} \right] \quad (294)$$

となる。 $q_\theta = \pi$ ならすべての候補が受理され、独立な標本が得られる。 q_θ が不完全でも、MH 補正後の定常分布は変わらず、近似の不足は主として受容率の低下や自己相関の増加として現れる。可逆な学習写像を proposal に使う場合はヤコビアンが提案確率比へ入り、学習した力や変換を HMC 軌道へ使う場合は軌道端の正確なハミルトニアン差で補正する。方式は異なっても、学習モデルが候補を作り、解析的に計算した受容規則が厳密性を回復するという構造は共通している。

ニューラルネットの近似には必ず誤差がある。 $q_\theta \simeq \pi$ であっても、MH 補正を行わず q_θ からのサンプル平均をそのまま物理結果とするとバイアスが残る。格子計算では、学習器は効率を改善し、最終的な定常分布は解析的な受容規則で保証する、という分業が基本になる。

ここでは3種類の誤差を区別する必要がある。第1は q_θ と π の差であり、これは提案の受容率や自己相関を悪化させる。第2は有限長のマルコフ連鎖による統計誤差である。第3は熱化不足や分布の一部へ到達できないことによる系統的な偏りである。Metropolis–Hastings 補正が除くのは、第1の差を物理量のバイアスへ持ち込むことであって、有限標本誤差やエルゴード性の問題まで消すわけではない。

したがって「厳密」という語は、補正後の遷移核が π を定常分布として持つという意味で用いる。実際の計算結果には、分布の被覆、熱化、自己相関、観測量の一致を別途検証する必要がある。

A33 リーフログ積分

実際には刻み幅 ϵ で数値積分する。1 ステップのリーフログ積分は

$$p_{n+1/2} = p_n - \frac{\epsilon}{2} \nabla S(\phi_n), \quad (295)$$

$$\phi_{n+1} = \phi_n + \epsilon M^{-1} p_{n+1/2}, \quad (296)$$

$$p_{n+1} = p_{n+1/2} - \frac{\epsilon}{2} \nabla S(\phi_{n+1}). \quad (297)$$

この積分器は可逆かつシンプレクティックで、位相空間体積を保存する。ただし有限 ϵ では $\Delta H_{\text{HMC}} = H_{\text{HMC}}(\phi', p') - H_{\text{HMC}}(\phi, p) \neq 0$ である。

そこで軌道末端を

$$A = \min(1, e^{-\Delta H_{\text{HMC}}}) \quad (298)$$

で受容する。数値積分誤差があっても、この Metropolis 補正によって目標分布は厳密に保たれる。

A34 シンプレクティック積分とエネルギー誤差

リーフログ積分は2次精度の積分器で、1 ステップの局所誤差は $O(\epsilon^3)$ 、一定長の軌道全体の誤差は $O(\epsilon^2)$ である。さらにシンプレクティック積分器には、HMC ハミルトニアン H_{HMC} に近いシャドー・ハミルトニアン

$$\tilde{H}_{\text{HMC}} = H_{\text{HMC}} + \epsilon^2 H_2 + \epsilon^4 H_4 + \cdots \quad (299)$$

を長時間ほぼ保存するという性質がある。そのため通常の非シンプレクティック積分器のようにエネルギー誤差が軌道長とともに1方向へ単調に蓄積しにくく、 ΔH_{HMC} は有界な振動として現れやすい。

ただし保存されるのは $\tilde{H}_{\text{HMC}}$ であって、目標分布を定める H_{HMC} そのものではない。最後の Metropolis 判定は、この H_{HMC} と $\tilde{H}_{\text{HMC}}$ の差を確率的に補正する。刻み幅を小さくすれば受容率は上がるが、力計算回数も増えるので、実用上は受容率と計算量の積を調整する必要がある。

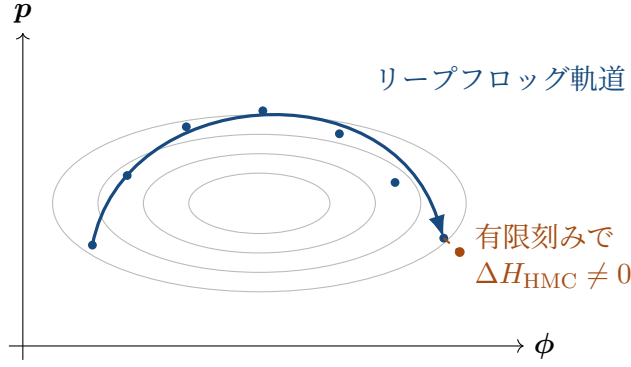


図 21: HMC 法の仮想位相空間。 ϕ と p は全格子自由度からなる高次元ベクトルであり、図では模式的に 2 次元へ投影している。厳密なハミルトン力学は等 H_{HMC} 面に沿い、受理時には軌道端の配位全体を大域的に更新する。数値積分によるずれ ΔH_{HMC} は最後の Metropolis 判定で補正する。

数値積分と Metropolis 補正の役割

リープフロッグ積分は、作用の勾配を使って場の全成分を同時に動かし、高い受容率を保ちながら配位空間を遠くへ移動する大域的提案を作る。一方、Metropolis 補正は、有限刻み幅によるハミルトニアン誤差を受容確率に反映し、目標分布を厳密に保つ。リープフロッグ積分だけでも、Metropolis 補正だけでも十分ではない。

別の積分器や学習した写像を使う場合にも、提案写像の可逆性と体積変化を明示しなければならない。体積を保存しない写像ではヤコビアン行列式が受容比に加わる。その受容比まで正しく構成して初めて、近似的な提案を厳密なマルコフ更新へ変換できる。

A35 2変数スカラー場に対する HMC 法の 1 更新

HMC を抽象的なアルゴリズムとして終わらせないため、最小の結合系を考える。周期境界条件を持つ 2 サイトの実スカラー場 $x = (x_1, x_2)$ に対し

$$S(x_1, x_2) = \frac{\kappa}{2}(x_1 - x_2)^2 + \frac{m^2}{2}(x_1^2 + x_2^2) + \frac{\lambda}{4}(x_1^4 + x_2^4) \quad (300)$$

とする。勾配は

$$\frac{\partial S}{\partial x_1} = \kappa(x_1 - x_2) + m^2 x_1 + \lambda x_1^3, \quad (301)$$

$$\frac{\partial S}{\partial x_2} = \kappa(x_2 - x_1) + m^2 x_2 + \lambda x_2^3. \quad (302)$$

隣接結合は x_1 と x_2 を近づけ、質量項は原点へ、4 次項は大振幅を強く抑える。

具体的に

$$\kappa = 1, \quad m^2 = 1, \quad \lambda = 1, \quad x = (1, 0), \quad p = (0, 1), \quad \epsilon = 0.1 \quad (303)$$

とする。初期作用と運動エネルギーは

$$S(1, 0) = \frac{1}{2} + \frac{1}{2} + \frac{1}{4} = 1.25, \quad K(p) = \frac{1}{2}, \quad H_{\text{HMC}}^{\text{old}} = 1.75. \quad (304)$$

初期勾配は

$$\nabla S(1, 0) = (3, -1) \quad (305)$$

なので、最初の半ステップは

$$p_{1/2} = p_0 - \frac{\epsilon}{2} \nabla S(x_0) \quad (306)$$

$$= (0, 1) - 0.05(3, -1) = (-0.15, 1.05). \quad (307)$$

位置の全ステップは

$$x_1 = x_0 + \epsilon p_{1/2} = (1, 0) + 0.1(-0.15, 1.05) = (0.985, 0.105). \quad (308)$$

新しい点での勾配は概算で

$$\nabla S(x_1) \simeq (2.821, -0.876) \quad (309)$$

であり、最後の半ステップは

$$p_1 = p_{1/2} - 0.05 \nabla S(x_1) \simeq (-0.291, 1.094). \quad (310)$$

この (x_1, p_1) から $H_{\text{HMC}}^{\text{new}}$ を計算し、 $e^{-\Delta H_{\text{HMC}}}$ で受容する。丸め誤差を除けば、 ϵ を小さくすると ΔH_{HMC} は急速に小さくなる。

A35.1 2変数模型から分かること

1. 勾配は確率の高い方向を直接教えるわけではなく、作用を下げる力を与える。
2. 運動量があるため、勾配降下のようにモードへ落ちて止まらず、典型集合を通過できる。
3. リープフロク積分は決定論的だが、軌道ごとに運動量を再生成するのでマルコフ連鎖は確率的になる。
4. 有限刻みの誤差は最後の MH test で補正される。

A35.2 ガウス分布における HMC 軌道

$\lambda = 0$ なら $S = \frac{1}{2} x^T K x$ であり、 $M = K$ を選ぶと固有モードごとの運動方程式は

$$\frac{d^2 x}{d\tau^2} = -x \quad (311)$$

となる。解は

$$x(\tau) = x(0) \cos \tau + K^{-1} p(0) \sin \tau. \quad (312)$$

HMC 法はガウス分布の等確率面に沿う回転として理解できる。質量行列 M を相関構造に合わせる前処理は、異なるモードの時間尺度を揃える操作である。

A36 ブラケットの連続極限

非可換ゲージ理論の連続極限で交換子がどこから来るかを見る。簡単のためリンクの始点を x とし

$$U_\mu(x) = e^{iagA_\mu(x)}, \quad (313)$$

$$U_\nu(x + a\hat{\mu}) = e^{iagA_\nu(x+a\hat{\mu})}, \quad (314)$$

$$U_\mu^\dagger(x + a\hat{\nu}) = e^{-iagA_\mu(x+a\hat{\nu})}, \quad (315)$$

$$U_\nu^\dagger(x) = e^{-iagA_\nu(x)} \quad (316)$$

を掛ける。隣接点の場を

$$A_\nu(x + a\hat{\mu}) = A_\nu(x) + a\partial_\mu A_\nu(x) + O(a^2), \quad (317)$$

$$A_\mu(x + a\hat{\nu}) = A_\mu(x) + a\partial_\nu A_\mu(x) + O(a^2) \quad (318)$$

と展開する。

Baker–Campbell–Hausdorff 公式

$$e^X e^Y = \exp \left(X + Y + \frac{1}{2}[X, Y] + \cdots \right) \quad (319)$$

を順に使う。 $O(a)$ の A_μ, A_ν は閉路を 1 周すると相殺する。 $O(a^2)$ では位置ずれによる微分項

$$iga^2(\partial_\mu A_\nu - \partial_\nu A_\mu) \quad (320)$$

と、行列が可換でないことによる

$$-g^2 a^2 [A_\mu, A_\nu] \quad (321)$$

が残る。まとめると

$$U_{\mu\nu}(x) = \exp \{ iga^2 (\partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu]) + O(a^3) \}. \quad (322)$$

括弧内が $F_{\mu\nu}$ である。

A36.1 可換な U(1) 理論との比較

U(1) では A_μ は数なので $[A_\mu, A_\nu] = 0$ である。従って

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (323)$$

SU(3) では $A_\mu = A_\mu^a T^a$ が行列であり、交換子がゲージ場自身の相互作用を生む。光子が電荷を持たないのに対し、gluon が色荷を担うことの作用レベルでの表現である。

A36.2 トレースと連続極限

$X = iga^2 F_{\mu\nu}$ として

$$\text{Re Tr } e^X = \text{Re Tr} \left(1 + X + \frac{1}{2}X^2 + \cdots \right). \quad (324)$$

$F_{\mu\nu}$ は Lie 代数に値を取り $\text{Tr } F_{\mu\nu} = 0$ なので 1 次項は消える。また $i^2 = -1$ から

$$\text{Re Tr } U_{\mu\nu} = N_c - \frac{g^2 a^4}{2} \text{Tr } F_{\mu\nu}^2 + O(a^6). \quad (325)$$

このため Wilson 作用の $1 - \text{Re Tr } U_p / N_c$ が正の F^2 へ戻る。

注意 ここで示したのは古典展開である。量子論として連続極限を取るには、裸の g_0 を格子間隔とともに running させる必要がある。QCD の漸近自由性により $a \rightarrow 0$ で $g_0 \rightarrow 0$ となり、この展開と整合する。

A37 Polyakov loop

有限温度では、時間方向を 1 周する Polyakov loop を

$$L(\mathbf{x}) = \frac{1}{N_c} \text{Tr} \prod_{t=0}^{N_t-1} U_4(\mathbf{x}, t) \quad (326)$$

で定義する。

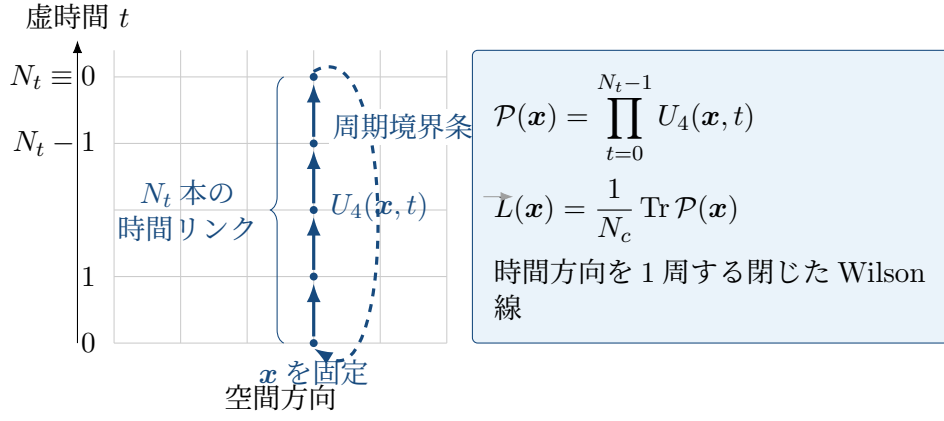


図 22: Polyakov loop の模式図。固定した空間点 x で時間方向のリンクを順に掛け、周期境界条件で始点へ戻る。積は始点で共役変換を受けるため、そのトレースは通常の周期的ゲージ変換に対して不変である。

Polyakov loop は中心対称性と閉じ込めの指標になる。

A38 アフィンカップリング層とヤコビアン計算

本文では、一般の $N \times N$ ヤコビアン行列式を計算すると $O(N^3)$ の計算量が必要になると述べた。アフィンカップリング層は、変換を意図的に block 三角構造へ制限し、可逆性と表現力を保ちながら対数ヤコビアンを $O(N)$ で評価する代表的な方法である。

入力を 2 組 $x = (x_A, x_B)$ に分け、 x_A からニューラルネット s_θ, t_θ を計算して

$$\begin{aligned} y_A &= x_A, \\ y_B &= x_B \odot \exp[s_\theta(x_A)] + t_\theta(x_A) \end{aligned} \quad (327)$$

と変換する。ここで $\odot$ は成分ごとの積である。 x_A を変えないため逆変換は

$$\begin{aligned} x_A &= y_A, \\ x_B &= [y_B - t_\theta(y_A)] \odot \exp[-s_\theta(y_A)] \end{aligned} \quad (328)$$

と陽に計算できる。 s_θ, t_θ を与えるニューラルネット自身は可逆である必要がない。

この変換のヤコビアンは

$$\frac{\partial(y_A, y_B)}{\partial(x_A, x_B)} = \begin{pmatrix} I & 0 \\ * & \text{diag}(e^{s_\theta(x_A)}) \end{pmatrix} \quad (329)$$

という block 三角行列になる。左下の $*$ にはニューラルネットの微分が現れるが、行列式には寄与しない。従って

$$\log \left| \det \frac{\partial y}{\partial x} \right| = \sum_j s_{\theta,j}(x_A) \quad (330)$$

であり、密行列の行列式を作らずに成分和だけで計算できる。ヤコビアン項を失わずに $O(N^3)$ から $O(N)$ へ下げられることが、この構成の中心である。

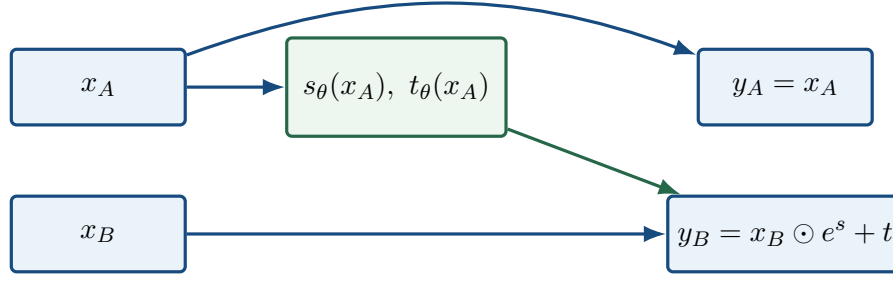


図 23: アフィンカップリング層。 x_A を固定したまま、その情報を使って x_B を伸縮・平行移動する。

1 層だけでは x_A が変化しないため、次の層では A と B を入れ替えるか、成分を並べ替える。この操作を繰り返すと、各層のヤコビアンは安価なまま、すべての成分が互いに影響する変換を作れる。格子場理論では、分割の仕方や s_θ, t_θ のネットワークに並進対称性、ゲージ共変性、周期境界条件などを組み込むことが、表現力だけでなく学習効率にも重要になる。

A39 1次元フローと MH 補正

A39.1 線形なスケール変換

基底分布を標準ガウス分布

$$r(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (331)$$

とし、 $x = az + b$ ($a > 0$) とする。逆変換は $z = (x - b)/a$ 、ヤコビアンは $dz/dx = 1/a$ なので

$$q(x) = \frac{1}{a\sqrt{2\pi}} \exp \left[-\frac{(x - b)^2}{2a^2} \right]. \quad (332)$$

フローは平均 b 、標準偏差 a のガウス分布を作った。線形変換だけでは多峰分布を作れないが、ヤコビアンの役割は完全に見える。

A39.2 非線形変換と確率密度

単調写像

$$x = f(z) = z + \alpha \tanh z, \quad \alpha > -1 \quad (333)$$

なら

$$\frac{dx}{dz} = 1 + \alpha \operatorname{sech}^2 z > 0 \quad (334)$$

で可逆である。密度は

$$\log q(f(z)) = -\frac{z^2}{2} - \frac{1}{2} \log(2\pi) - \log(1 + \alpha \operatorname{sech}^2 z). \quad (335)$$

$\alpha > 0$ なら原点付近をより強く引き伸ばし、原点付近の密度を薄くする。ニューラルネットは、この局所的な伸縮を高次元で組み合わせる。

A39.3 不完全な提案分布に対する MH 補正

目標を 2 峰分布

$$\pi(x) \propto \frac{1}{2}e^{-(x-2)^2/2} + \frac{1}{2}e^{-(x+2)^2/2} \quad (336)$$

とする。学習した $q(x)$ が右側のモードを多く、左側を少なく生成するとする。そのまま q の平均を取ればバイアスする。

独立 MH では現在点 x と候補 x' に対し

$$R = \frac{\pi(x')q(x)}{\pi(x)q(x')} \quad (337)$$

を計算する。 q が過剰に出す領域では $q(x')$ が分母で大きくなり、受容率が下がる。 q が過少な領域から出る方向では逆に補償される。この比が提案頻度の偏りを正す。

A39.4 確率分布の台に関する条件

もし $q(x) = 0$ だが $\pi(x) > 0$ となる領域があれば、フローはそこを決して提案できない。MH 補正は「出てきた候補の重み」を直せるが、「絶対に出ない候補」を作れない。従って厳密性には

$$\text{supp}(\pi) \subseteq \text{supp}(q) \quad (338)$$

が必要である。ガウス基底分布と全空間上の可逆写像はこの点で扱いやすい。

A39.5 受容率と重みのばらつき

重要度重み

$$w(x) = \frac{e^{-S(x)}}{q_\theta(x)} \quad (339)$$

を考えると、独立 MH の比は $R = w(x')/w(x)$ である。 $q_\theta = \pi$ なら w は定数で全提案を受容する。 w の分散が大きいと、稀な大きい重みの状態に入った連鎖が長く滞在し、自己相関が増える。従って平均受容率だけでなく、 $\log w$ の分布や effective sample size を見る必要がある。

A40 モード崩壊と分布の被覆

逆 KL は q が確率を置いた場所でのみ π とのずれを罰する。従って多峰分布では、1 つのモードだけを覆っても損失が小さくなりうる。トポロジカルセクターの取りこぼしは、局所的な観測量や高い受容率だけでは見抜けない場合がある。

田中と富谷は、restricted Boltzmann machine と HMC 法を組み合わせ、3 次元 ϕ^4 理論で自己相関を短縮する初期の機械学習例を調べた [22]。一部の観測量では通常の HMC 法と整合した一方、2 点関数や臨界点近傍の分布には差も残った。この結果は、自己相関や作用密度だけでなく、相関関数と分布の被覆を検証する必要性を示している。

検証には少なくとも次が必要である。

1. 作用、プラケット、相関関数など複数の観測量。
2. トポロジカルチャージの履歴と分布。
3. 独立な初期条件や複数の連鎖間の一致。
4. 厳密性が確認されたサンプラーまたは既知結果との比較。

A41 有限時間幅と有効質量

数値計算では Euclid 時間方向の長さ $T = N_t a$ は有限であり、通常は周期境界条件を課す。このとき粒子が $t = 0$ から t へ進む寄与に加え、周期境界をまたいで $T - t$ だけ進む寄与がある。最低状態だけを残すと

$$C(t) \simeq A \left(e^{-mt} + e^{-m(T-t)} \right) \quad (340)$$

$$= 2Ae^{-mT/2} \cosh \left[m \left(\frac{T}{2} - t \right) \right] \quad (341)$$

となる。したがって単純な指数関数を時間中央 $t \simeq T/2$ まで用いるのは正しくない。

隣接する 3 点から定義する cosh 型有効質量は

$$m_{\text{eff}}(t) = \frac{1}{a} \operatorname{arcosh} \left[\frac{C(t-a) + C(t+a)}{2C(t)} \right] \quad (342)$$

である。式 (341) が 1 状態で支配される領域では $m_{\text{eff}}(t)$ が一定値を取る。実際の解析では、この plateau の位置を変え、励起状態の混入と統計誤差の双方を調べる。

より一般には

$$C(t) = \sum_n A_n \left(e^{-\Delta E_n t} + e^{-\Delta E_n (T-t)} \right), \quad \Delta E_n = E_n - E_0. \quad (343)$$

短い t では多数の励起状態が寄与し、長い t では信号自体が指数的に小さくなる。質量抽出は、単に大きい t を取ればよい問題ではなく、励起状態と雑音の競合を扱う問題である。

A42 重要公式一覧

中心となる関係式をまとめる。

$$\langle O \rangle = Z^{-1} \int \mathcal{D}\phi \, O[\phi] e^{-S[\phi]}, \quad (344)$$

$$U_{\mu\nu}(n) = U_\mu(n) U_\nu(n + \hat{\mu}) U_\mu^\dagger(n + \hat{\nu}) U_\nu^\dagger(n), \quad (345)$$

$$S_g = \beta \sum_{n, \mu < \nu} \left(1 - \frac{1}{3} \operatorname{Re} \operatorname{Tr} U_{\mu\nu}(n) \right), \quad (346)$$

$$Z_{\text{QCD}} = \int \mathcal{D}U \, e^{-S_g} \prod_f \det D_f[U], \quad (347)$$

$$H_{\text{HMC}}(\phi, p) = S(\phi) + \frac{1}{2} p^2, \quad (348)$$

$$A_{\text{HMC}} = \min(1, e^{-\Delta H_{\text{HMC}}}), \quad (349)$$

$$q_\theta(x) = r(z) \left| \det \frac{\partial z}{\partial x} \right|, \quad (350)$$

$$A_{\text{flow}} = \min \left[1, \frac{e^{-S(x')} q_\theta(x)}{e^{-S(x)} q_\theta(x')} \right]. \quad (351)$$

参考文献

- [1] K. Osterwalder and R. Schrader, “Axioms for Euclidean Green’s Functions,” Commun. Math. Phys. **31**, 83–112 (1973), doi:10.1007/BF01645738.

- [2] A. S. Wightman, “Quantum Field Theory in Terms of Vacuum Expectation Values,” *Phys. Rev.* **101**, 860–866 (1956), doi:10.1103/PhysRev.101.860.
- [3] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.* **21**, 1087–1092 (1953), doi:10.1063/1.1699114.
- [4] W. K. Hastings, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika* **57**, 97–109 (1970), doi:10.1093/biomet/57.1.97.
- [5] J. Schwinger, “On Gauge Invariance and Vacuum Polarization,” *Phys. Rev.* **82**, 664–679 (1951), doi:10.1103/PhysRev.82.664.
- [6] C. N. Yang and R. L. Mills, “Conservation of Isotopic Spin and Isotopic Gauge Invariance,” *Phys. Rev.* **96**, 191–195 (1954), doi:10.1103/PhysRev.96.191.
- [7] R. Utiyama, “Invariant Theoretical Interpretation of Interaction,” *Phys. Rev.* **101**, 1597–1607 (1956), doi:10.1103/PhysRev.101.1597.
- [8] K. G. Wilson, “Confinement of quarks,” *Phys. Rev. D* **10**, 2445–2459 (1974), doi:10.1103/PhysRevD.10.2445.
- [9] K. G. Wilson, “Quarks and Strings on a Lattice,” in *New Phenomena in Subnuclear Physics, Part A*, Proceedings of the 1975 International School of Subnuclear Physics, Erice, Sicily, edited by A. Zichichi (Plenum Press, New York, 1977), pp. 69–142, report CLNS-321.
- [10] S. Elitzur, “Impossibility of Spontaneously Breaking Local Symmetries,” *Phys. Rev. D* **12**, 3978–3982 (1975), doi:10.1103/PhysRevD.12.3978.
- [11] M. Gell-Mann and F. E. Low, “Quantum Electrodynamics at Small Distances,” *Phys. Rev.* **95**, 1300–1312 (1954), doi:10.1103/PhysRev.95.1300.
- [12] J. Polchinski, “Renormalization and Effective Lagrangians,” *Nucl. Phys. B* **231**, 269–295 (1984), doi:10.1016/0550-3213(84)90287-6.
- [13] K. G. Wilson and J. Kogut, “The Renormalization Group and the ϵ Expansion,” *Phys. Rep.* **12**, 75–199 (1974), doi:10.1016/0370-1573(74)90023-4.
- [14] K. Symanzik, “Continuum Limit and Improved Action in Lattice Theories. I. Principles and ϕ^4 Theory,” *Nucl. Phys. B* **226**, 187–204 (1983), doi:10.1016/0550-3213(83)90468-6.
- [15] M. Creutz, Monte Carlo Study of Quantized SU(2) Gauge Theory, *Phys. Rev. D* **21**, 2308–2315 (1980), doi:10.1103/PhysRevD.21.2308.
- [16] H. B. Nielsen and M. Ninomiya, “No Go Theorem for Regularizing Chiral Fermions,” *Phys. Lett. B* **105**, 219 (1981).
- [17] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, “Hybrid Monte Carlo,” *Phys. Lett. B* **195**, 216 (1987).
- [18] J. Liu, Y. Qi, Z. Y. Meng and L. Fu, “Self-Learning Monte Carlo Method,” *Phys. Rev. B* **95**, 041101(R) (2017), doi:10.1103/PhysRevB.95.041101, arXiv:1610.03137 [cond-mat.str-el].
- [19] G. Cybenko, “Approximation by Superpositions of a Sigmoidal Function,” *Math. Control Signals Systems* **2**, 303–314 (1989), doi:10.1007/BF02551274.
- [20] H. Nicolai, “Supersymmetry and Functional Integration Measures,” *Nucl. Phys. B* **176**, 419–428 (1980), doi:10.1016/0550-3213(80)90460-5.

- [21] M. Lüscher, Trivializing Maps, the Wilson Flow and the HMC Algorithm, *Commun. Math. Phys.* **293**, 899–919 (2010), arXiv:0907.5491.
- [22] A. Tanaka and A. Tomiya, “Towards Reduction of Autocorrelation in HMC by Machine Learning,” arXiv:1712.03893 [hep-lat] (2017).
- [23] S. Foreman, T. Izubuchi, L. Jin, X.-Y. Jin, J. C. Osborn and A. Tomiya, “HMC with Normalizing Flows,” *PoS LATTICE2021*, 073 (2022), doi:10.22323/1.396.0073, arXiv:2112.01586.
- [24] P. Boyle, T. Izubuchi, L. Jin, C. Jung, C. Lehner, N. Matsumoto and A. Tomiya, “Use of Schwinger–Dyson Equation in Constructing an Approximate Trivializing Map,” *PoS LATTICE2022*, 229 (2023), doi:10.22323/1.430.0229, arXiv:2212.11387.
- [25] D. J. Rezende and S. Mohamed, Variational Inference with Normalizing Flows, *Proc. 32nd Int. Conf. Machine Learning*, PMLR **37**, 1530–1538 (2015), arXiv:1505.05770.
- [26] L. Dinh, J. Sohl-Dickstein and S. Bengio, Density Estimation Using Real NVP, *International Conference on Learning Representations* (2017), arXiv:1605.08803.
- [27] J. Ho, A. Jain and P. Abbeel, “Denoising Diffusion Probabilistic Models,” *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020), arXiv:2006.11239.
- [28] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon and B. Poole, “Score-Based Generative Modeling through Stochastic Differential Equations,” *International Conference on Learning Representations* (2021), arXiv:2011.13456.
- [29] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui and M.-H. Yang, “Diffusion Models: A Comprehensive Survey of Methods and Applications,” *ACM Computing Surveys* (2023), doi:10.1145/3626235, arXiv:2209.00796 [cs.LG].
- [30] L. Wang, G. Aarts and K. Zhou, “Diffusion Models as Stochastic Quantization in Lattice Field Theory,” *JHEP* **05**, 060 (2024), arXiv:2309.17082 [hep-lat], doi:10.1007/JHEP05(2024)060.
- [31] Q. Zhu, G. Aarts, W. Wang, K. Zhou and L. Wang, “Physics-Conditioned Diffusion Models for Lattice Gauge Theory,” *JHEP* **03**, 111 (2026), arXiv:2502.05504 [hep-lat], doi:10.1007/JHEP03(2026)111.
- [32] A. Tomiya, “Lattice Configuration Generation with a Self-Learning Diffusion Model,” arXiv:2607.12587 [hep-lat] (2026), doi:10.48550/arXiv.2607.12587.
- [33] M. S. Albergo, G. Kanwar and P. E. Shanahan, “Flow-based generative models for Markov chain Monte Carlo in lattice field theory,” *Phys. Rev. D* **100**, 034515 (2019), arXiv:1904.12072.
- [34] Y. Nagai and A. Tomiya, “Gauge covariant neural network for quarks and gluons,” *Phys. Rev. D* **111**, 074501 (2025), doi:10.1103/PhysRevD.111.074501, arXiv:2103.11965.
- [35] M. Favoni, A. Ipp, D. I. Müller and D. Schuh, “Lattice gauge equivariant convolutional neural networks,” *Phys. Rev. Lett.* **128**, 032003 (2022), doi:10.1103/PhysRevLett.128.032003, arXiv:2012.12901.
- [36] Y. Nagai, H. Ohno and A. Tomiya, “CASK: A Gauge Covariant Transformer for Lattice Gauge Theory,” *PoS LATTICE2024*, 030 (2025), doi:10.22323/1.466.0030, arXiv:2501.16955.
- [37] G. Kanwar et al., “Equivariant flow-based sampling for lattice gauge theory,” *Phys. Rev. Lett.* **125**, 121601 (2020), arXiv:2003.06413.
- [38] D. Boyda et al., “Sampling using $SU(N)$ gauge equivariant flows,” *Phys. Rev. D* **103**, 074504 (2021), arXiv:2008.05456.

- [39] R. Abbott et al., Gauge-Equivariant Flow Models for Sampling in Lattice Field Theories with Pseudofermions, *Phys. Rev. D* **106**, 074506 (2022), doi:10.1103/PhysRevD.106.074506.
- [40] R. Abbott et al., “Sampling QCD field configurations with gauge-equivariant flow models,” PoS LATTICE2022, 036 (2023), arXiv:2208.03832 [hep-lat].
- [41] R. Abbott et al., “Progress in Normalizing Flows for 4d Gauge Theories,” arXiv:2502.00263 [hep-lat] (2025).
- [42] Y. Nagai and A. Tomiya, “JuliaQCD: Portable lattice QCD package in Julia language,” arXiv:2409.03030 [hep-lat] (2024).
- [43] C. Gattringer and C. B. Lang, *Quantum Chromodynamics on the Lattice: An Introductory Presentation*, Lecture Notes in Physics **788** (Springer, 2010), doi:10.1007/978-3-642-01850-3.
- [44] J. Bezanson, A. Edelman, S. Karpinski and V. B. Shah, “Julia: A Fresh Approach to Numerical Computing,” *SIAM Rev.* **59**, 65–98 (2017), doi:10.1137/141000671.
- [45] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall/CRC, 1993), doi:10.1201/9780429246593.
- [46] A. Tomiya, “JkTools: A toolkit for Jackknife method,” GitHub repository, commit fa8f234 (2026), <https://github.com/akio-tomiya/JkTools>.
- [47] M. Guagnelli, R. Sommer and H. Wittig, “Precision computation of a low-energy reference scale in quenched lattice QCD,” *Nucl. Phys. B* **535**, 389–402 (1998), doi:10.1016/S0550-3213(98)00599-9, arXiv:hep-lat/9806005.