Introduction to Lattice Field Theory

Sourendu Gupta

TIFR

Asian School on Lattice Field Theory
TIFR, Mumbai, India
March 12, 13, 2011

Outline

The quantum problem

A quantum problem with Hamiltonian H is completely specified if one can compute the unitary evolution operator

$$U(0,T) = \mathrm{e}^{i\int_0^T dt H(t)}$$

There are path integral representations for this operator. All our study starts from here.
The finite temperature quantum problem is completely understood if one computes the partition function

$$Z(\beta) = \mathrm{Tr}\,\mathrm{e}^{-\int_0^\beta dt H(t)},$$

which is formally the same problem in Euclidean time. The same path integral suffices to solve this problem.

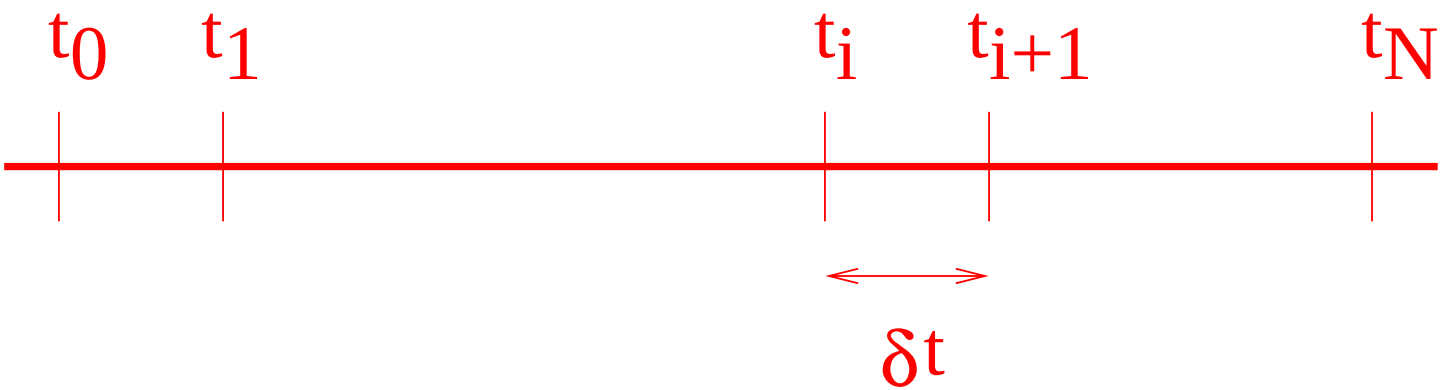
A path integral is matrix multiplication

Define $\delta t = T/N_t$ The amplitude for a quantum state $|x_0\rangle$ at initial time 0 to evolve to the state $|x_N\rangle$ at the final time T can be written as

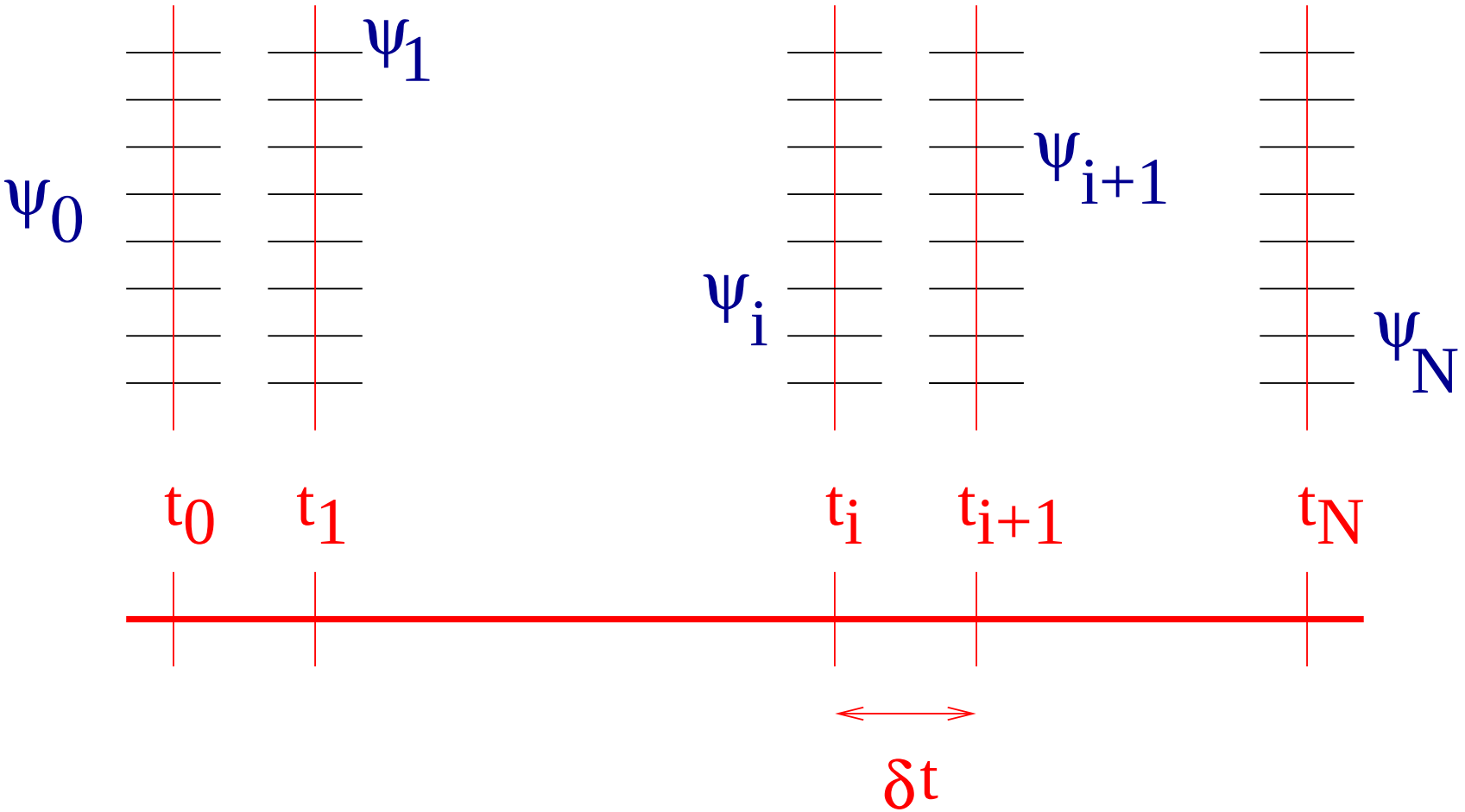
$$\begin{aligned}\langle \alpha_N | \, U(0,T) \, |\alpha_0\rangle \quad = \quad & \sum_{\psi_1,\psi_2,\cdots,\psi_{N-1}} \langle \alpha_N | \, U((N-1)\delta t, T) \, |\psi_{N-1}\rangle \\ & \langle \psi_{N-1} | \, U(\delta t, (N-1)\delta t) \, |\psi_{N-2}\rangle \cdots \\ & \langle \psi_1 | \, U(\delta t, 0) \, |\alpha_0\rangle \, ,\end{aligned}$$

where we have inserted complete sets of states at the end of each interval. The notation also distinguishes between the states at the end points and the basis states $|\psi_i\rangle$ at the intermediate points. This sum over all intermediate states is called the **path integral**. The choice of the basis states $|\psi_i\rangle$ is up to us, and we can choose them at our convenience.

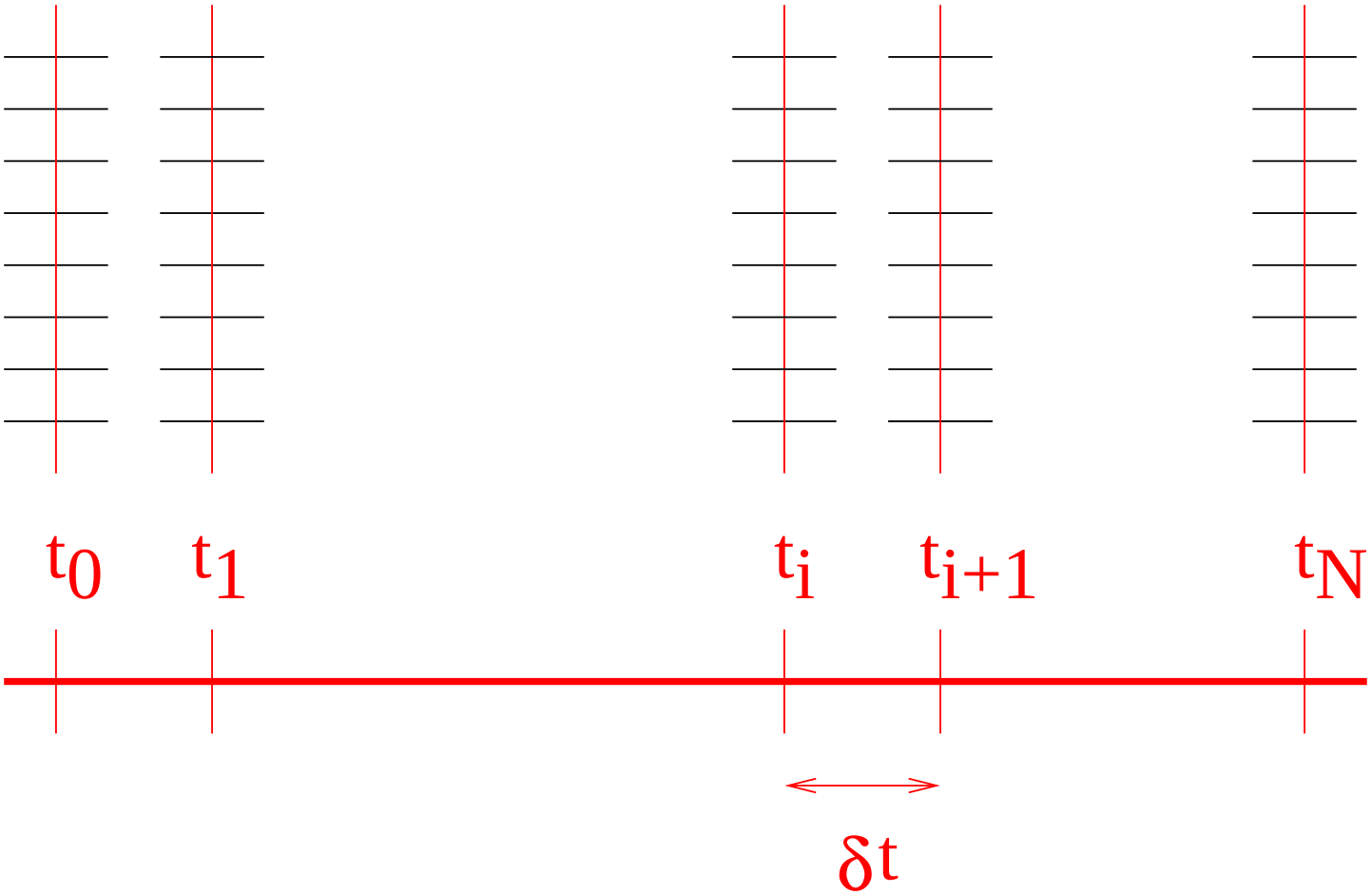
A path integral



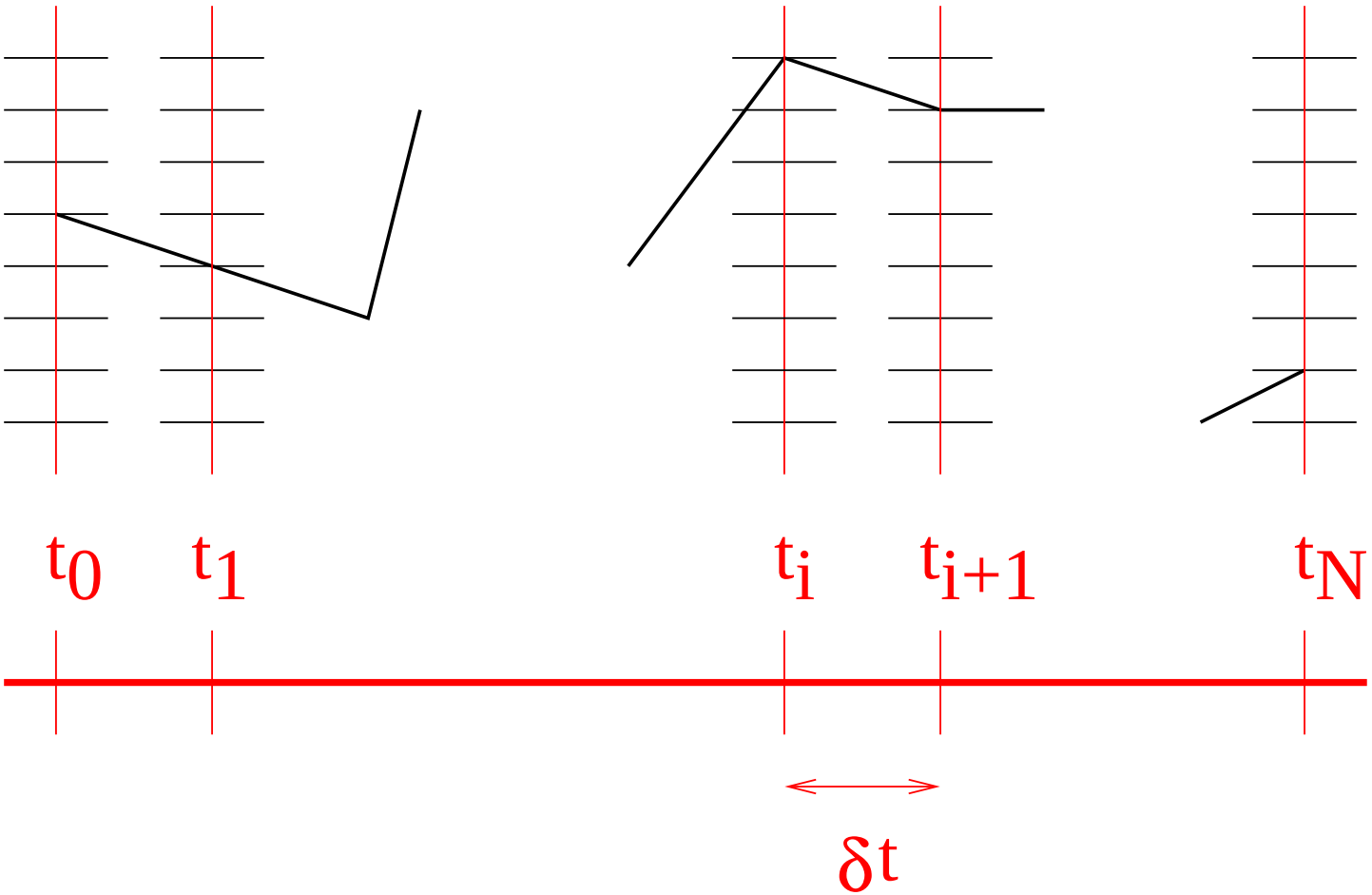
A path integral



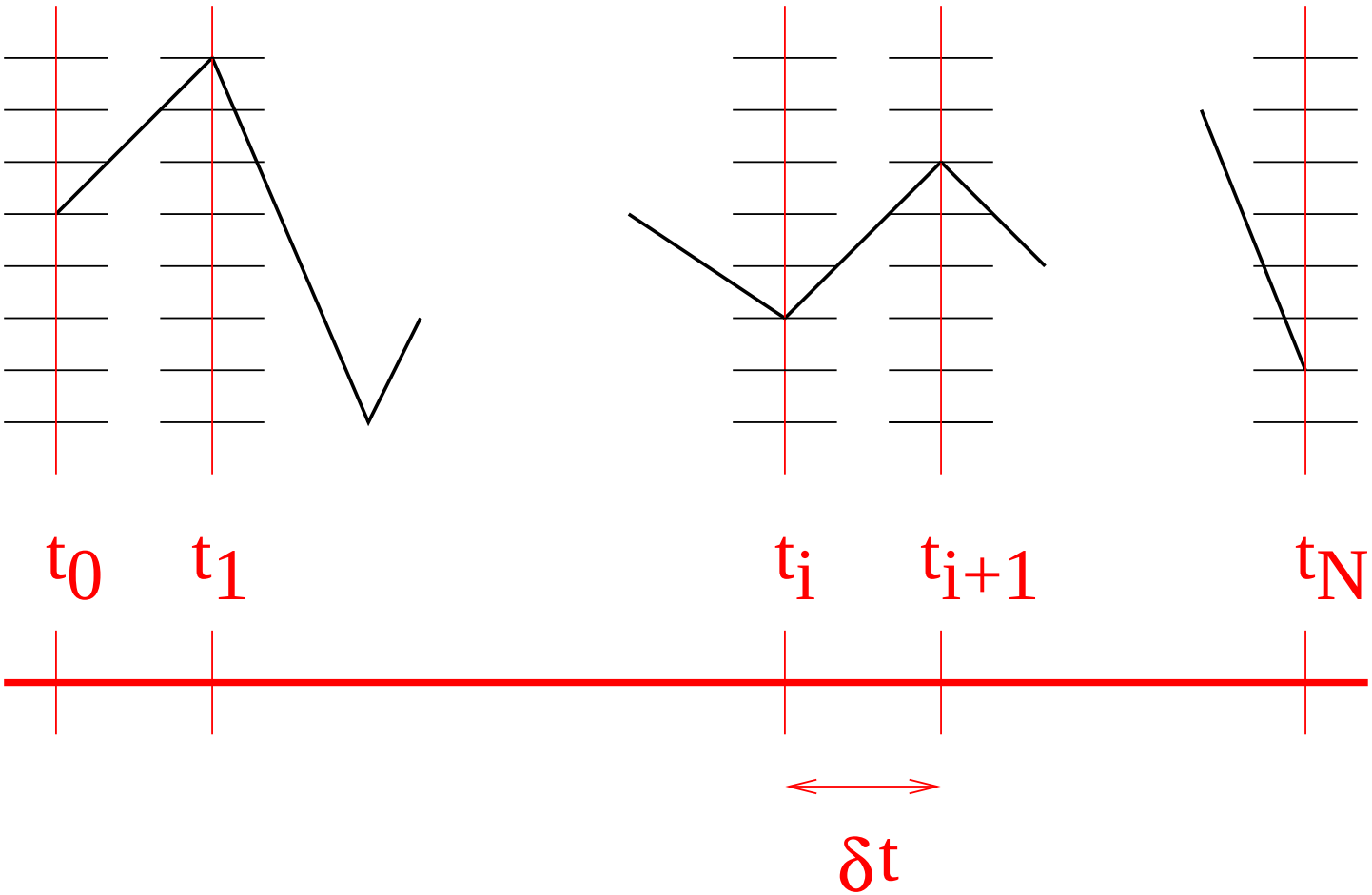
A path integral



A path integral



A path integral



Choosing the basis states

If the basis states $|\psi_i\rangle$ are eigenstates of the Hamiltonian then

$$\langle\psi_{i+1}|\, U(t_i+\delta t, t_i)\,|\psi_i\rangle = \mathrm{e}^{-iE_i\delta t/\hbar}\delta_{\psi_{i+1},\psi_i}.$$

The result looks trivial because the hard task of diagonalizing the Hamiltonian is already done. In any other basis the path integral is non-trivial.
By choosing position eigenstates as the basis, Feynman [1] showed that the infinitesimal evolution operator for a single particle is given by

$$\langle x|\, U(0,\delta t)\,|y\rangle = \mathrm{e}^{i\delta t S(x,y)},$$

where $S(x,y)$ is the classical action for a trajectory of the particle which goes from the point x at time 0 to the point y at time δt.
For a spin problem one may use angular momentum coherent states, for fermions Grassman coherent states. The results are similar.

Going to the diagonal basis

If V is unitary and $V^\dagger H V$ is diagonal, then

$$U(0,T) = V^\dagger \begin{pmatrix} \mathrm{e}^{-iE_0 T} & 0 & \cdots \\ 0 & \mathrm{e}^{-iE_1 T} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} V.$$

The sum over intermediate states is diagonal, and the Vs act only on the initial and final states to give

$$\begin{aligned} \langle \alpha_N | \, U(0,T) \, | \alpha_0 \rangle &= (\alpha_N^0)^* \alpha_0^0 \mathrm{e}^{-iE_0 T} + (\alpha_N^1)^* \alpha_0^1 \mathrm{e}^{-iE_1 T} + \cdots \\ &\longrightarrow (\alpha_N^0)^* \alpha_0^0 \mathrm{e}^{-E_0 T}, \quad (\text{Euclidean time } t \to -it), \end{aligned}$$

when $E_1 > E_0$ and $T(E_1 - E_0) \gg 1$. This gives the lowest eigenvalue of H.

Introducing the transfer matrix

The Euclidean problem over one lattice step in time is now phrased in terms of the **transfer matrix** —

$$T(\delta t) = U(0, -i\delta t) = V^\dagger \begin{pmatrix} \mathrm{e}^{-E_0\delta t} & 0 & 0 & \dots \\ 0 & \mathrm{e}^{-E_1\delta T} & 0 & \dots \\ 0 & 0 & \mathrm{e}^{-E_2\delta T} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} V.$$

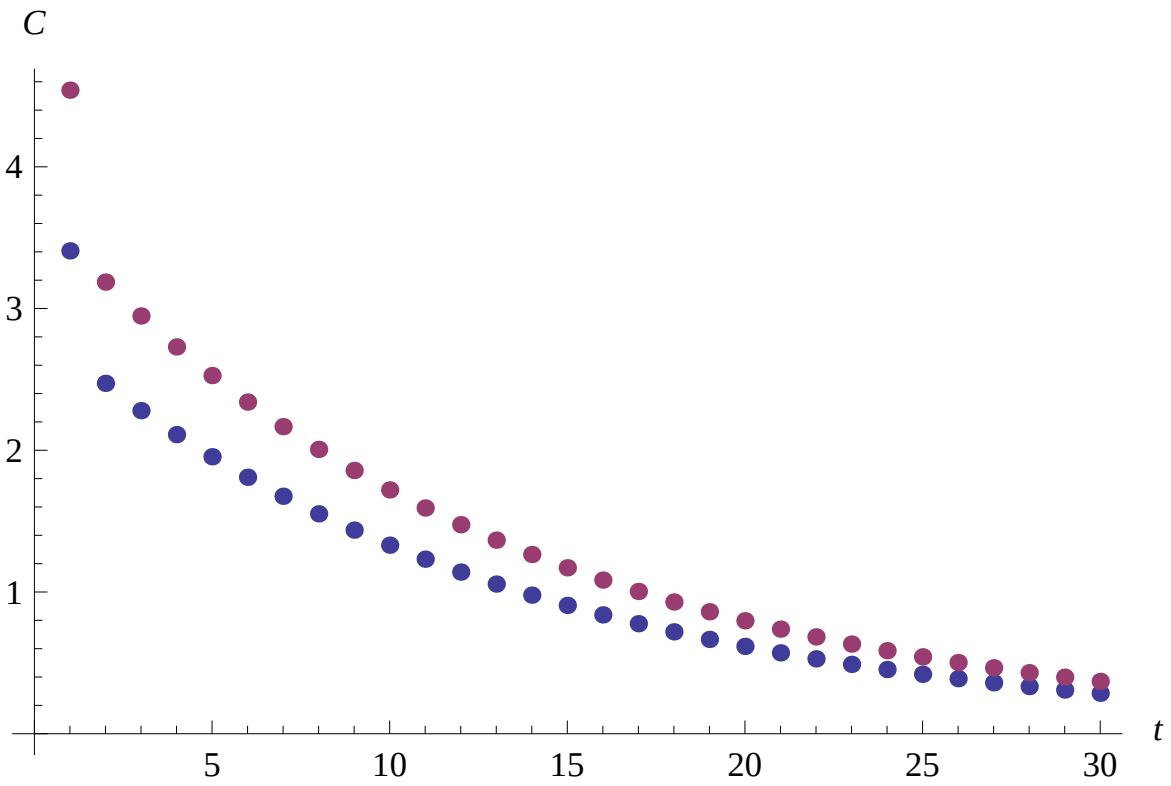
Since $T = \exp[-\delta t H]$, the two operators commute, and have the same eigenvectors. If the eigenvalues of T are called λ_i, then the eigenvalues of the Hamiltonian are $E_i = -(\log \lambda_i)/\delta t$.
If we are to recover a finite E_i when $\delta t \to 0$, then $\log \lambda$ must go to zero. The **correlation length** in the problem is $\xi = 1/\log \lambda$, so this must diverge in order to give finite E_i. Therefore the **continuum limit** corresponds to a **critical point**.

Algorithm for computing energies

For a new formulation of quantum mechanics we have a trivial algorithm for computing the energy. It exploits the simple fact that given a randomly chosen unit vector $|\phi\rangle$, the matrix element $\langle\phi| T^n |\phi\rangle$ tends to λ_0^n as $n \to \infty$.

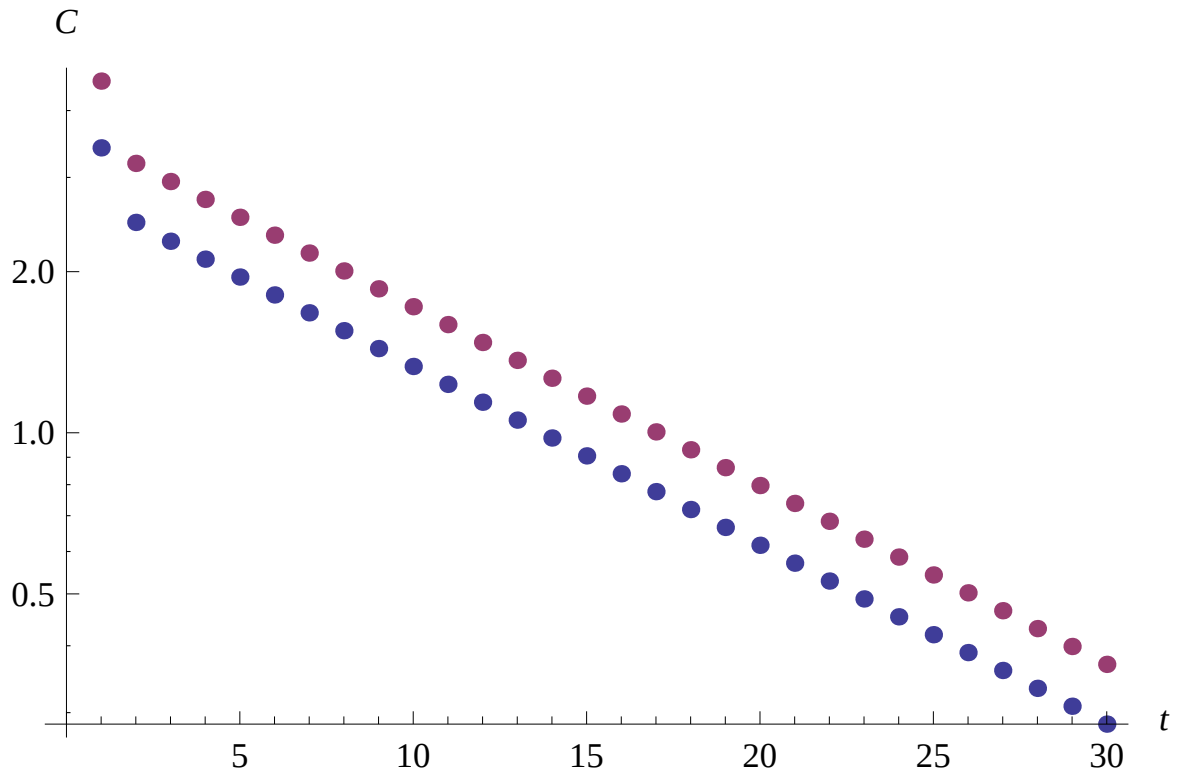
1. Choose a **source**. At one time slice construct a random linear combination of basis states: $|\phi_0\rangle$.
2. Choose a path **configuration**, *i.e.*, a random $|\phi_j\rangle$ on each lattice site $(j\delta t)$ with probability **given by** T. Construct a measurement of the **correlation function** $C_{0j} = \langle\phi_j\phi_k\rangle$.
3. Repeat step 2 as many times as feasible and construct the mean $\langle C_{0j}\rangle = \langle\phi_0| T^j |\phi_0\rangle$ (since $|\phi_j\rangle$ are chosen with appropriate weight, the mean suffices).
4. Plot $log\,\langle C_{0j}\rangle$ against j. At sufficiently large j the slope gives $-E_0\delta t$. Alternatively, find a plateau in the **local masses** $m_j = \log(\langle C_{0,j+1}\rangle / \langle C_{0j}\rangle)$.

A measurement of a correlation function



Correlation functions decrease monotonically because all eigenvalues of T are positive (**reflection positivity**) as a consequence of the unitarity of U.

A measurement of a correlation function



Correlation functions decrease monotonically because all eigenvalues of T are positive (**reflection positivity**) as a consequence of the unitarity of U.

Quantum field theory

Quantum mechanics of a single particle is a 1-dimensional field theory. The (Euclidean) Feynman path integral is

$$Z = \int \mathcal{D}\mathbf{x} \exp\left[-\int_{-\infty}^{\infty} dt\ S(\mathbf{x})\right],$$

where S is the action and the integral over an $\mathbf{x}$ at each time is **regularized** by discretizing time.
We extend this to a quantum field theory in dimension D. If the space-time points are labelled by x, and the fields are $\phi(x)$, then the Euclidean partition function is

$$Z = \int \mathcal{D}\phi \exp\left[-\int d^D x\ S(\phi)\right],$$

where S is the action density and the integrals may be regulated by discretizing space-time.

The lattice and the reciprocal lattice

In the usual perturbative approach to field theory, the computation of any n-point function involves loop integrals which diverge. These are regulated by putting a cutoff Λ on the 4-momentum. When space-time is regulated by discretization, then the lattice spacing a provides the cutoff $\Lambda = 1/a$.
We will take the discretization of space-time to be a regular hypercubic lattice, with sites denoted by a vector of integers $x = a\mathbf{j}$. When fields are placed on such a lattice, $\phi(x)$, the momenta are no longer continuous, but form a reciprocal lattice.

$$k = \frac{2\pi}{a}\mathbf{l},$$

where the $\mathbf{l}$ are integers. The physics at all points on the reciprocal lattice are equivalent.

Fourier transforms and the Brillouin zone

In practice our lattice will not be infinite, but a finite hypercube with, say, N^D sites. At each site, x, on the lattice, let us put a complex number $\phi(x)$. We can put periodic boundary conditions, $\phi(x) = \phi(x + \hat{\mu}N)$. Next, we can make Fourier transforms

$$\phi(k) = \sum_x \phi(x) \mathrm{e}^{ik\cdot x}, \quad \phi(x) = \frac{1}{N^D} \sum_k \phi(k) \mathrm{e}^{-ik\cdot x},$$

where $k = 2\pi \mathbf{l}/N$, and $\mathbf{l}$ have components taking integer values between 0 and N or $-N/2$ and $N/2$. One sees that the boundary condition allows components only inside the **Brillouin zone**, *i.e.*, the region $-\pi/a \leq k_\mu \leq \pi/a$. Any k outside this is mapped back inside by the periodicity of the lattice.
The completeness of the Fourier basis implies

$$\frac{1}{N^D} \sum_x \mathrm{e}^{-iq\cdot x} = \delta_{0q\cdot}$$

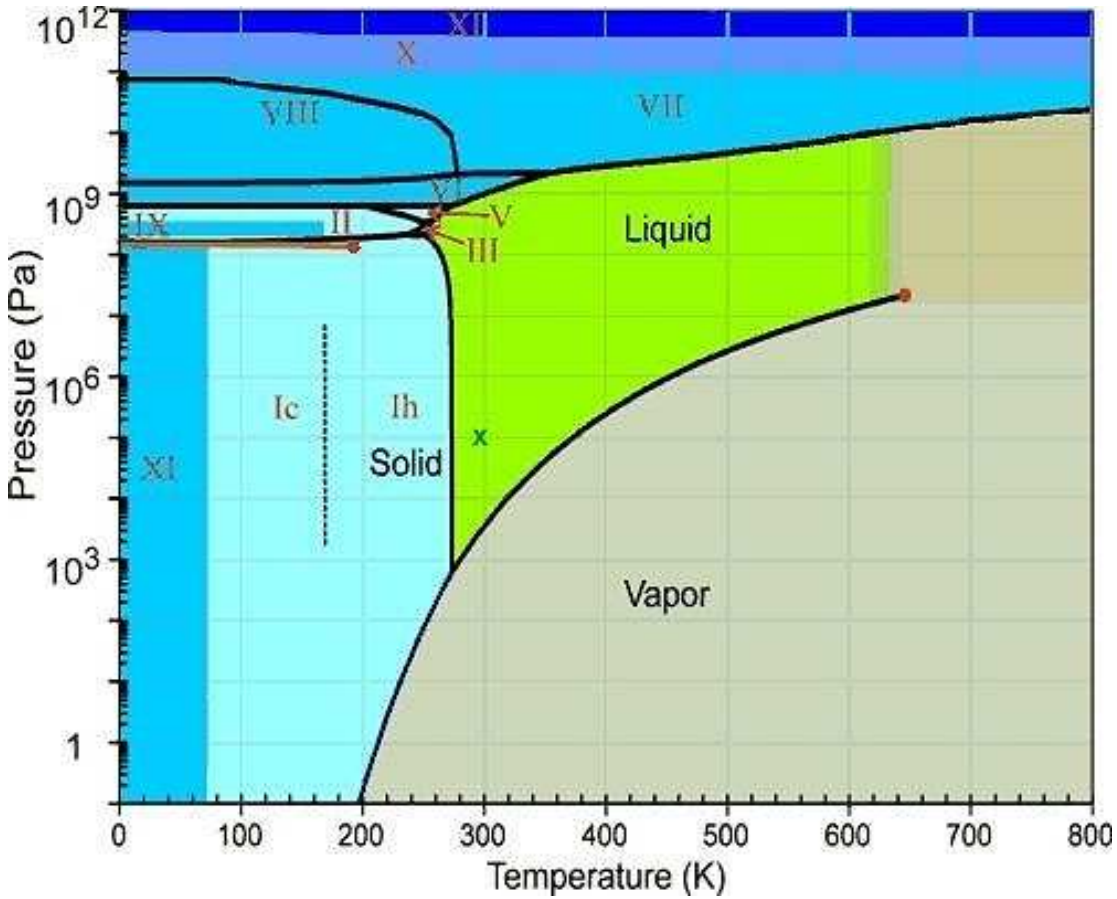
The great unification

The renormalization procedure will be to take the continuum limit $a \to 0$ (*i.e.*, $\Lambda \to \infty$) keeping some physical quantity fixed, such as a mass, m. If this is fixed in physical units, then in lattice units it must diverge as $a \to 0$. This corresponds to a second order phase transition on the lattice.

A lattice field theory in Euclidean time and D dimensions of space is exactly the same as a statistical mechanics on a $D+1$ dimensional lattice. Here is the precise analogy—

Action	$\leftrightarrow$	Transfer matrix
Path integral	$\leftrightarrow$	Partition function
2-point function	$\leftrightarrow$	Correlation function
Continuum limit	$\leftrightarrow$	2nd order phase transition
Unitarity	$\leftrightarrow$	Reflection positivity

Phase transitions



Normal single phase behaviour, two-phase coexistence (first order phase transitions), three-phase coexistence (triple points), critical point (second order phase transitions).

Divergences and critical exponents

Scaling of free energy at a critical point (T_c, P_c)

$$F(T,P) = p^a f\left(\frac{t}{p^b}\right) \qquad p = P - P_c,\ t = T - T_c.$$

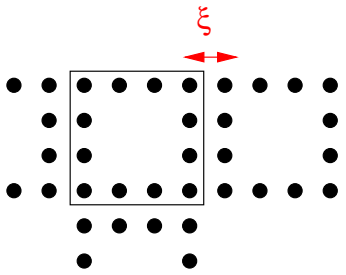
The scaling form implies power law divergence of the specific heat ($t^{-\alpha}$), order parameter ($t^{-\beta}$ and $p^{-\delta}$) and order parameter susceptibility ($t^{-\gamma}$) at the critical point. There are various relations between these **critical exponents** since the scaling form contains only two exponents (see [3]).
Also there is scaling of the correlation function—

$$G(r, T, p = 0) = r^{(2-d-\eta)} g\left(\frac{r}{t^{-\nu}}\right).$$

At the critical point the correlation length diverges. The scaling form implies that

$$\xi \propto t^{-\nu}, \qquad G(r, t = 0, p = 0) \propto \frac{1}{r^{\eta+d-2}}.$$

Coarse graining and the Renormalization Group



If the correlation length of a system is ξ, then one can try to define **coarse grained** variables by summing over blocks of sites. When the block size becomes larger than ξ, the problem simplifies.

A renormalization group (RG) transformation is the following—

1. Coarse grain by summing the field over a block of size ζa, and scale the sum to the same range as the original fields. This changes $a \to \zeta a$.
2. Find the Hamiltonian of the coarse grained field which reproduces the thermodynamics of the original system. The couplings in the Hamiltonians "flow" $g(a) \to g(\zeta a) = g'$.
3. The flow follows the Callan-Symanzik **beta-function**

$$B(g) = -\frac{\partial g}{\partial \zeta},$$

(note the sign). A **fixed point** of the RG has $B(g) = 0$.

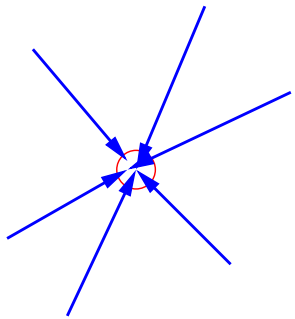
Linearized Renormalization Group transformation

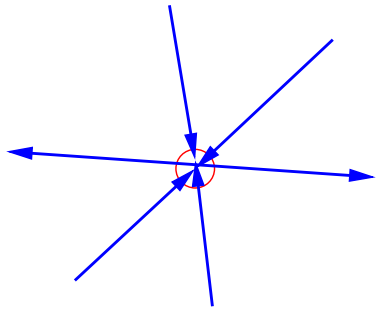
Assume that there are multiple couplings G_i with beta-functions B_i. At the critical point the values are G_i^c. Define $g_i = G_i - G_i^c$. Then,

$$B_i(G_1, G_2, \cdots) = \sum_j B_{ij} g_j + \mathcal{O}(g^2).$$

Diagonalize the matrix B whose elements are B_{ij}. In cases of interest the eigenvalues turn out to be real.
Eigenvectors corresponding to negative eigenvalues are called **relevant operators**, for positive eigenvalues, the eigenvectors are called **irrelevant operators**, those with zero eigenvalues are called **marginal operators**.
If an eigenvalue is y then the corresponding eigenvector $v \to \zeta^{-y} v$ under an RG scaling by factor ζ. The eigenvalues are called **anomalous dimensions**. Marginal operators correspond to logarithmic scaling.

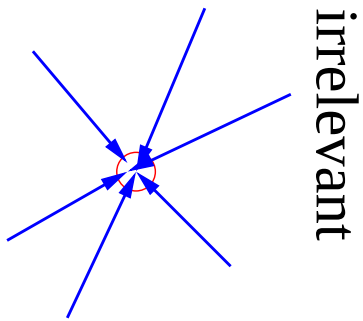
Renormalization Group trajectories

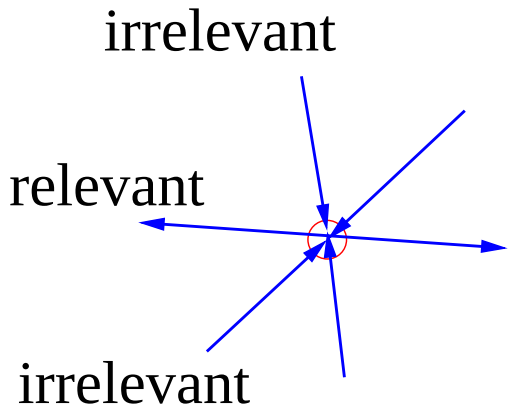




Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

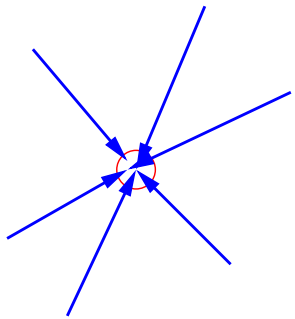
Renormalization Group trajectories

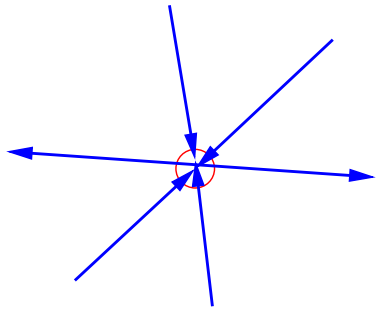




Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

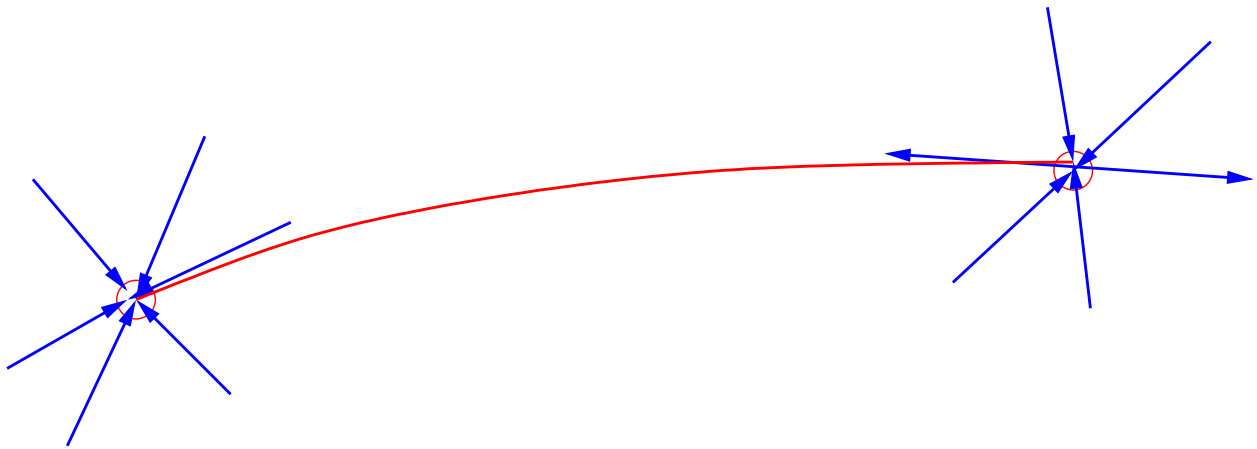
Renormalization Group trajectories





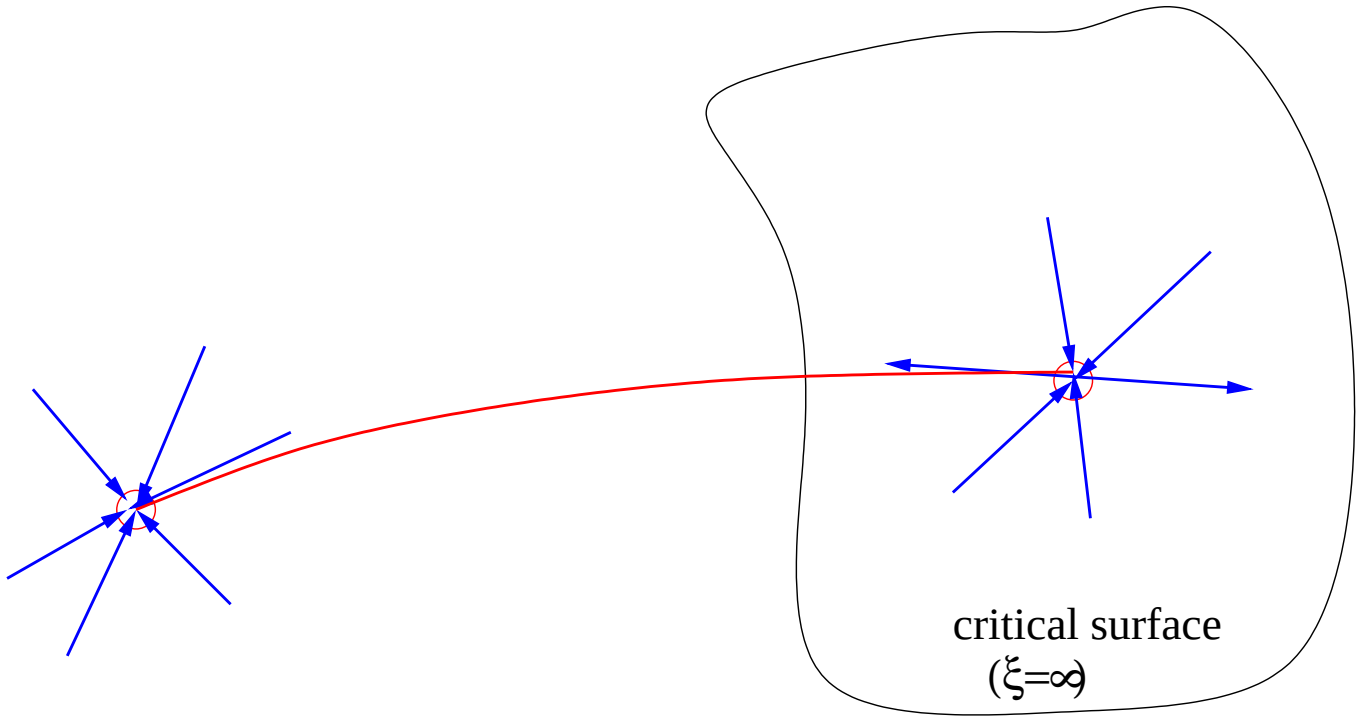
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



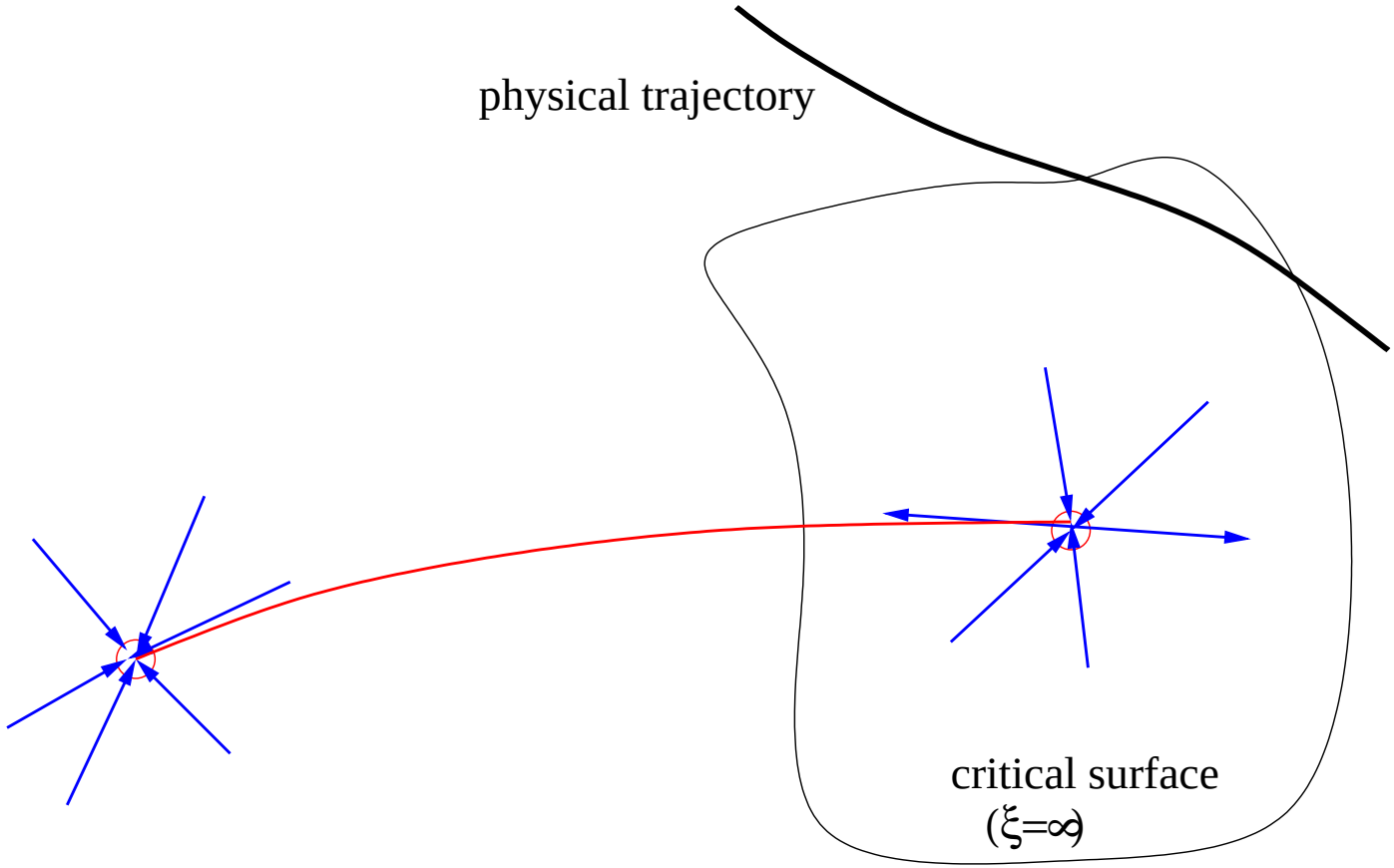
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



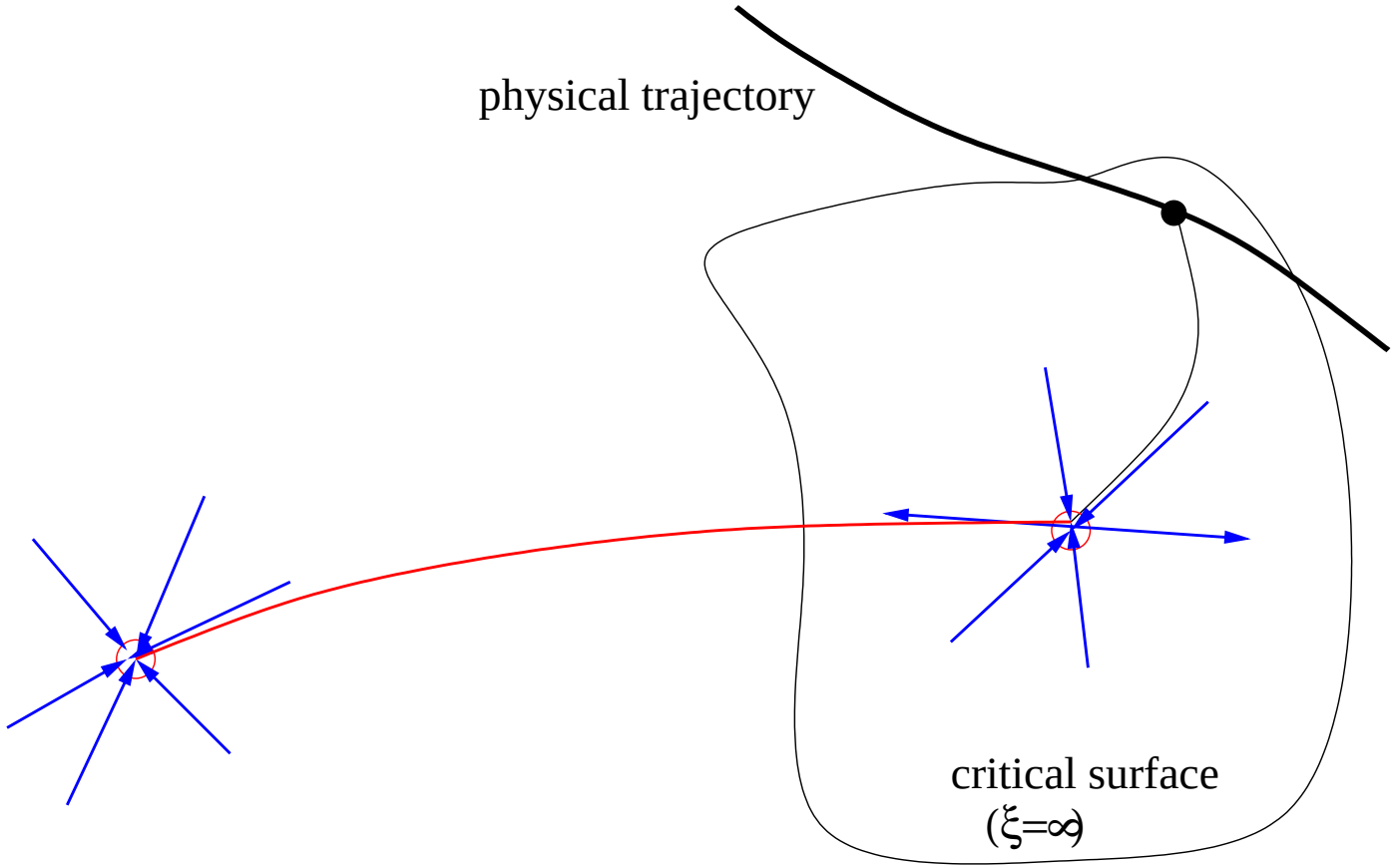
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



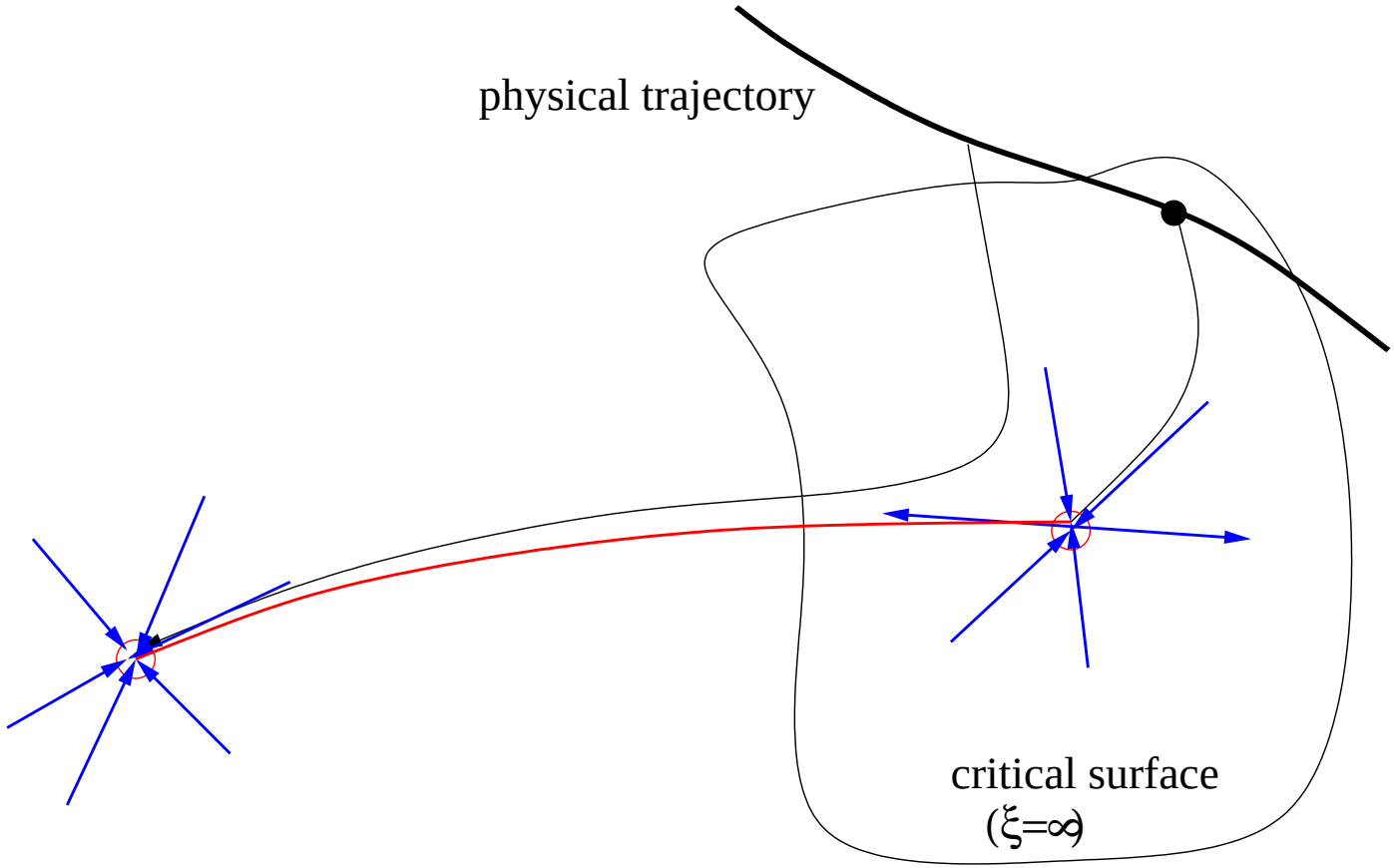
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



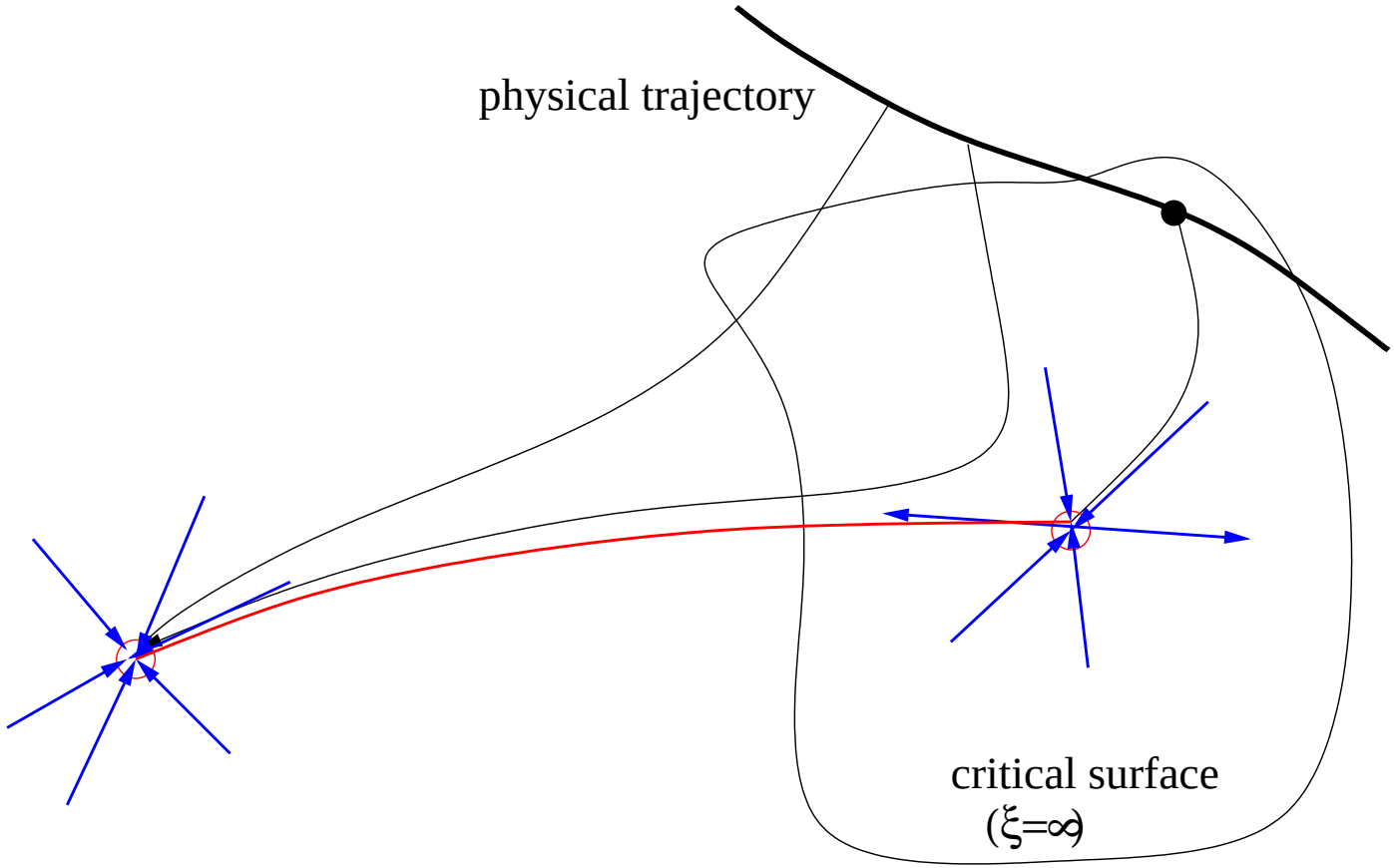
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



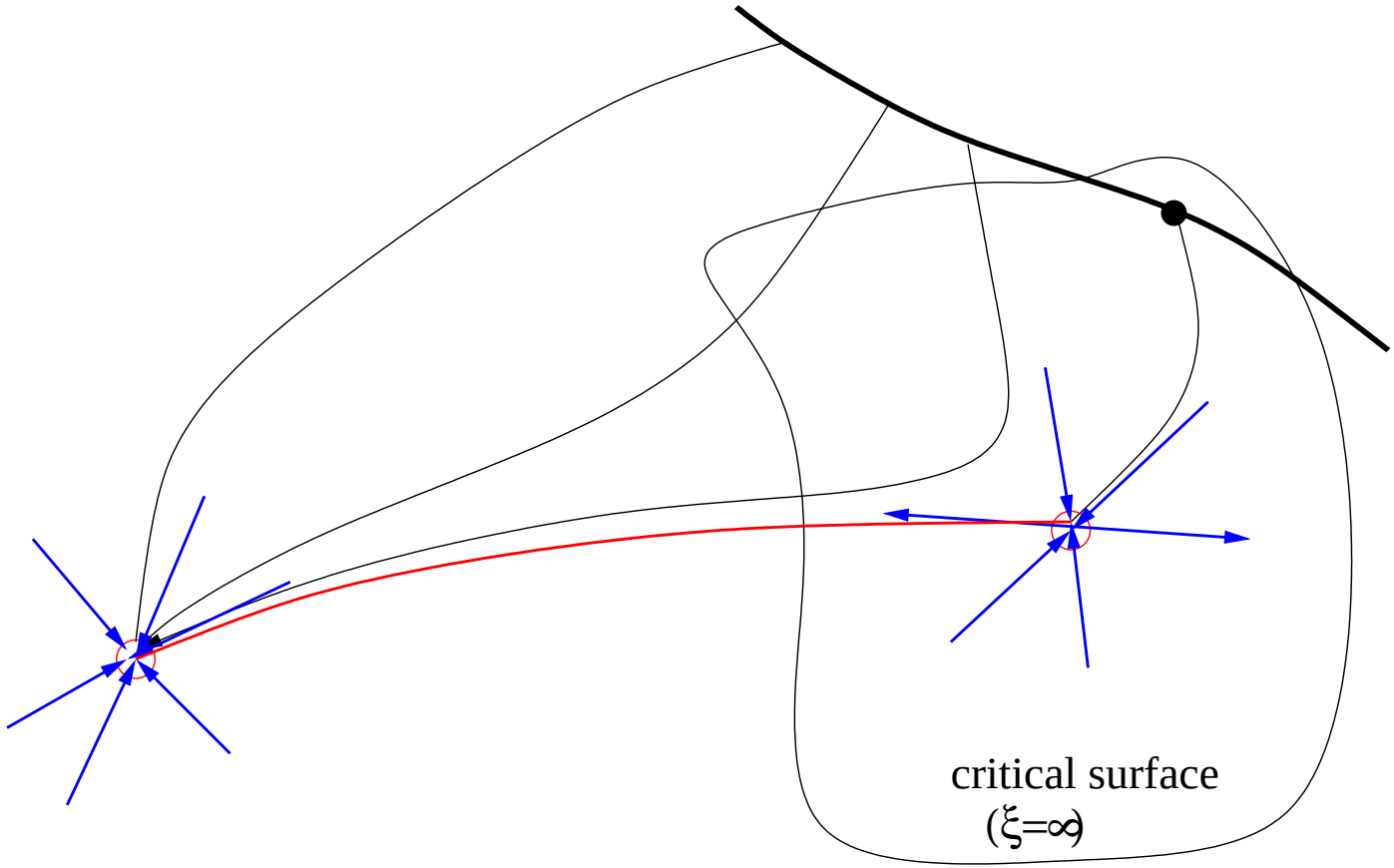
Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Renormalization Group trajectories



Fixed points: $\xi = 0$ (stable) or $\xi = \infty$ (unstable).

Probability theory as a trivial case of field theory

Generate random variables x with a probability distribution $P(x)$.
We can always shift our definition of x so that $\langle x \rangle = 0$.
It is useful to introduce the moment generating function

$$Z(j) = \sum_n \langle x^n \rangle \frac{j^n}{n!} = \int dx e^{xj} P(x).$$

The derivatives of $Z(j)$ give the moments. Now define the **characteristic function** $F(j) = \log Z(j)$. The derivatives give cumulants. We will use the notation

$$\left[x^2\right] = \left.\frac{d^2 F(j)}{dj^2}\right|_{j=0} = \left\langle x^2 \right\rangle - \langle x \rangle^2 = \sigma^2.$$

The Hamiltonian of statistical mechanics is analogous to $h(x) = \log P(x)$. Then $Z(j)$ is the partition function, and $F(j)$ the free energy. The derivatives of F give expectations of connected parts; these are the cumulants.

Coarse graining and the RG

Take a group of N random numbers, x_i, and define their mean

$$ {}_N x = \frac{1}{N} \sum_i x_i . $$

The ${}_N x$ are **coarse grained** random variables. A standard question in probability theory is the distribution of these coarse grained variables. Clearly this is a question in RG.
We need to compute the coarse grained characteristic function $F_N(j)$. First,

$$ \begin{aligned} Z_N(j) &= \int d_N x \, \mathrm{e}^{{}_N x j} \delta \left({}_N x - \frac{1}{N} \sum_{i=1}^{N} x_i \right) \prod_{i=1}^{N} dx_i \mathrm{e}^{h(x_i)} \\ &= [Z(j/N)]^N . \qquad \text{implies} \\ F_N(j) &= N F \left(\frac{j}{N} \right) . \end{aligned} $$

The central limit theorem: a fixed point theorem

Since

$$F(j) = \sigma^2 \frac{j^2}{2!} + [x^3] \frac{j^3}{3!} + [x^4] \frac{j^4}{4!} + \cdots ,$$

we find the RG flow gives

$$F_N(j) = \frac{\sigma^2}{N} \frac{j^2}{2!} + \frac{[x^3]}{N^2} \frac{j^3}{3!} + \frac{[x^4]}{N^3} \frac{j^4}{4!} + \cdots .$$

In the limit, since all the higher cumulants scale to zero much faster, we find that the RG flows to the $\mathrm{Gaussian\ fixed\ point}$ $F_N(j) = \sigma^2 j^2/(2N)$. This is the content of the **central limit theorem**.

Subtleties may occur if $\sigma^2 = 0$, with extensions to the case when all the cumulants up to some order are zero. Other subtleties arise when the distributions are fat-tailed and all the cumulants diverge. Other RG methods are needed for these special cases.

Generating functions in field theory

Any integral with a non-negative integrand can be treated as a $D = 0$ field theory. Some of the tricks one plays with integrals can be generalized to field theories.
In any field theory it is useful to extend the path integral to a generating functional of correlation functions—

$$Z[J] = \int \mathcal{D}\phi \exp\left[- \int d^D x \; S(\phi) + J(x)\phi(x) \right],$$

The connected parts of correlation functions are recovered as usual by taking functional derivatives—

$$C(z, z') = \frac{1}{Z[J]} \frac{\delta^2 Z[J]}{\delta J(z) \delta J(z')} \bigg|_{J=0}.$$

These are clear generalization of the notions of the moment generating function and the characteristic function.

The Ising model

The Ising model on a one-dimensional lattice contains a "spin" variable, $\sigma_i = \pm 1$ at each site, i, of the lattice. The Hamiltonian is

$$H = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1}.$$

We may put periodic boundary conditions on the lattice through the condition that $\sigma_{N+1} = \sigma_1$. We write $\beta = J/T$.
This can be solved by introducing the transfer matrix [2]

$$T(\beta) = \begin{pmatrix} \mathrm{e}^{\beta} & \mathrm{e}^{-\beta} \\ \mathrm{e}^{-\beta} & \mathrm{e}^{\beta} \end{pmatrix}.$$

Since $Z(\beta) = \mathrm{Tr}\, T^N$, the eigenvalues of the transfer matrix completely specify the solution. We find

$$Z(\beta) = 2^N \left[(\cosh \beta)^N + (\sinh \beta)^N \right].$$

The system becomes ordered only in the limit $\beta \to \infty$.

Coarse graining and fixed points

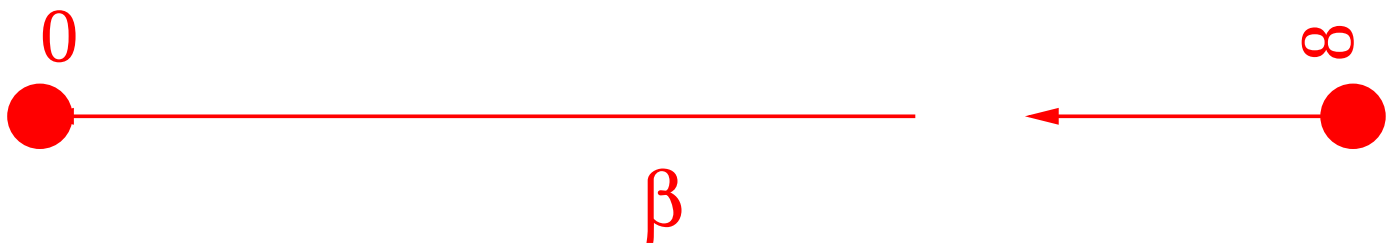
However, we can also perform a coarse graining with $\zeta = 2$. Since

$$T^2(\beta) = \begin{pmatrix} 2\cosh\beta & 2 \\ 2 & 2\cosh\beta \end{pmatrix} = \sqrt{2\cosh\beta} \begin{pmatrix} z & 1/z \\ 1/z & z \end{pmatrix},$$

where $z = \sqrt{(\cosh\beta)/2}$. Expanding around $\beta = \infty$ one may define the renormalized temperature as

$$\beta' = \beta - \frac{1}{2}\log 2 + \mathcal{O}\left(\mathrm{e}^{-4\beta}\right).$$

The fixed point is $\beta \to \infty$, as expected. This is a repulsive fixed point. At the other end, one has $\beta' = \beta^2 + \mathcal{O}(\beta^4)$. Hence $\beta = 0$ is another fixed point, which is attractive.



The RG flow for the 1-d Ising model is particularly simple.

Power counting

Consider the relativistic quantum field theory of a single real scalar field ϕ. The Lagrangian density, $\mathcal{L}$, can be written as a polynomial in the field and its derivatives. One usually encounters the terms

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}M^2\phi^2 + \frac{g_3}{3!}\phi^3 + \frac{g_4}{4!}\phi^4 + \cdots$$

Let us count the mass dimensions of the fields in units of a length L or a momentum Λ. Since the action is dimensionless, $[\mathcal{L}] = L^{-D} = \Lambda^D$. The kinetic term shows that

$$[\phi] = L^{1-D/2} = \Lambda^{D/2-1}.$$

The couplings have dimensions

$$\begin{aligned} \left[M^2\right] &= L^{-2} = \Lambda^2, \\ [g_3] &= L^{(D-6)/2} = \Lambda^{(6-D)/2}, \\ [g_4] &= L^{D-4} = \Lambda^{4-D}. \end{aligned}$$

The upper critical dimension

For each operator in the theory there is a certain dimension at which the coupling is marginal. This is called the **upper critical dimension**, D_u. The coupling g_r corresponding to the operator ϕ^r has

$$D_u = \frac{2r}{r-2}.$$

The mass is a relevant coupling in all dimensions, g_3 is relevant below $D_u = 6$, g_4 below $D_u = 4$. All other operators are irrelevant in $D = 4$. Derivative couplings are relevant (the kinetic term is marginal) in all dimensions.
Bogoliubov and Shirkov [5] set out power counting rules for divergences of loop integrals. It turns out that for $D > D_u$ an operator is unrenormalizable; at $D = D_c$ the operator gives a renormalizable contribution, and for $D < D_u$ the theory is super-renormalizable.

Field theory is not statistical mechanics

Divergences in statistical mechanics are due to long-distance physics. In field theory they are due to short distance physics. Therefore, in statistical mechanics it is the power of L which counts. For field theory, it is instead the power of Λ which determines which terms are important.
This is also reflected in the differences in the physical meaning of RG transformations in the two cases. The critical point in statistical mechanics is a point in the phase diagram where the correlation length actually becomes infinite. In field theory the critical point can be reached for any mass of the particle by scaling the lattice spacing to zero (momentum cutoff to infinity).

$$\begin{aligned} \text{irrelevant} &\leftrightarrow \text{un-renormalizable} \\ \text{marginal} &\leftrightarrow \text{renormalizable} \\ \text{relevant} &\leftrightarrow \text{super-renormalizable} \end{aligned}$$

Scalar Field Theory

The continuum Lagrangian for a single component real scalar field theory can be easily written for the lattice

$$\begin{aligned} S &= a^D \sum_x \frac{1}{2a^2} \sum_\mu [\phi(x+\hat{\mu}a) - \phi(x)]^2 + \frac{1}{2} m^2 \phi^2(x) + V(\phi) \\ &= \sum_x M^2 \phi^2(x) - \sum_\mu [\phi(x)\phi(x+\hat{\mu})] + V(\phi). \end{aligned}$$

In the first line we have replaced derivatives by the forward difference, ∇, on the lattice, and kept dimensional variables explicit. The notation is that x denotes a lattice site, μ one of the D directions, $\hat{\mu}$ an unit vector in that direction and a the lattice spacing. In the second line we have absorbed appropriate powers of a into every variable, written out the expressions in dimensionless units and then set $a = 1$. Note that $M^2 = D + m^2 a^2/2$.
By the earlier power-counting, it suffices to take $V(\phi) = g_4 \phi^4/4!$ in $D \geq 4$.

Notation for lattice theories

1. The lattice spacing will always be written as a except when we use units where $a = 1$.
2. We will use the notation x, y, *etc.*, to denote either a point in continuum space-time, or on the lattice.
3. Fourier transforms on N^D lattices are

$$\phi(k) = \sum_x \phi(x) \mathrm{e}^{ik\cdot x}, \quad \phi(x) = \frac{1}{N^D} \sum_k \phi(k) \mathrm{e}^{-ik\cdot x}, \quad k = 2\pi \mathbf{i}/N,$$

where reciprocal lattice points $\mathbf{i}$ have components taking values between 0 and N or $-N/2$ and $N/2$. In other words, the Brillouin zone contains momenta between $\pm\pi$. The completeness of the Fourier basis implies

$$\frac{1}{N^D} \sum_x \mathrm{e}^{-iq\cdot x} = \delta_{0q}.$$

Free scalar field theory

On a lattice of size N^D the free theory, $V = 0$, can be completely solved by Fourier transformation. The action becomes

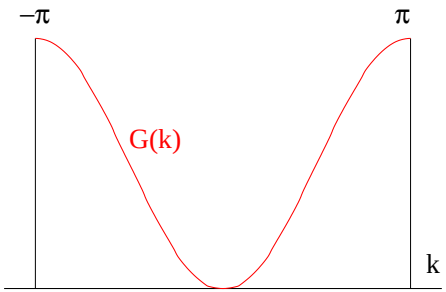
$$S = \sum_k \frac{1}{2} \left[m^2 - \sum_\mu (1 - \cos k_\mu) \right] \phi^2(k).$$

Since the Fourier transform is an unitary transformation of fields, the Jacobian for going from $\phi(x)$ to $\phi(k)$ is unity. Therefore, the Fourier transformation gives a set of decoupled Gaussian integrals, and

$$\begin{aligned} Z[\beta] &= \int \prod_k d\phi(k) \mathrm{e}^{-\beta S} = \prod_k \left[\frac{\beta}{2} (m^2 + \sum_\mu \sin^2 \frac{k_\mu}{2}) \right]^{-1/2} \\ &= \frac{1}{\sqrt{\det \frac{\beta}{2} (\nabla^2 + m^2)}}, \end{aligned}$$

where ∇ is the forward difference operator.

Low energy modes and Symanzik improvement



When $m = 0$ the two-point function of scalar field theory, G, vanishes at the center of the Brillouin zone and is maximum at the edges. Inside the Brillouin zone there is only one long distance mode when $a \to 0$.

At small k one has $G \simeq k^2[1 + \mathcal{O}(k^2 a^2)]$. **Symanzik improvement** consists of improving the a-dependence at tree level at finite lattice spacing by adding irrelevant terms to the lattice action. For a scalar field, one can write

$$G(k) = \frac{4}{3}(1 - \cos k_\mu) - \frac{1}{12}(1 - \cos 2k_\mu) = \frac{1}{2}k^2 + \mathcal{O}(k^6).$$

Hence, by removing the k^4 terms, one has an improved action. Clearly this is achieved by taking the forward difference and the two-step forward difference with appropriate coefficients.

The interacting theory

The standard form of the action for the scalar theory is

$$S = \sum_x \left[V(\phi) - \kappa \sum_\mu \phi(x)\phi(x+\hat{\mu}) \right], \quad V(\phi) = \lambda(\phi^2 - 1)^2 - \phi^2.$$

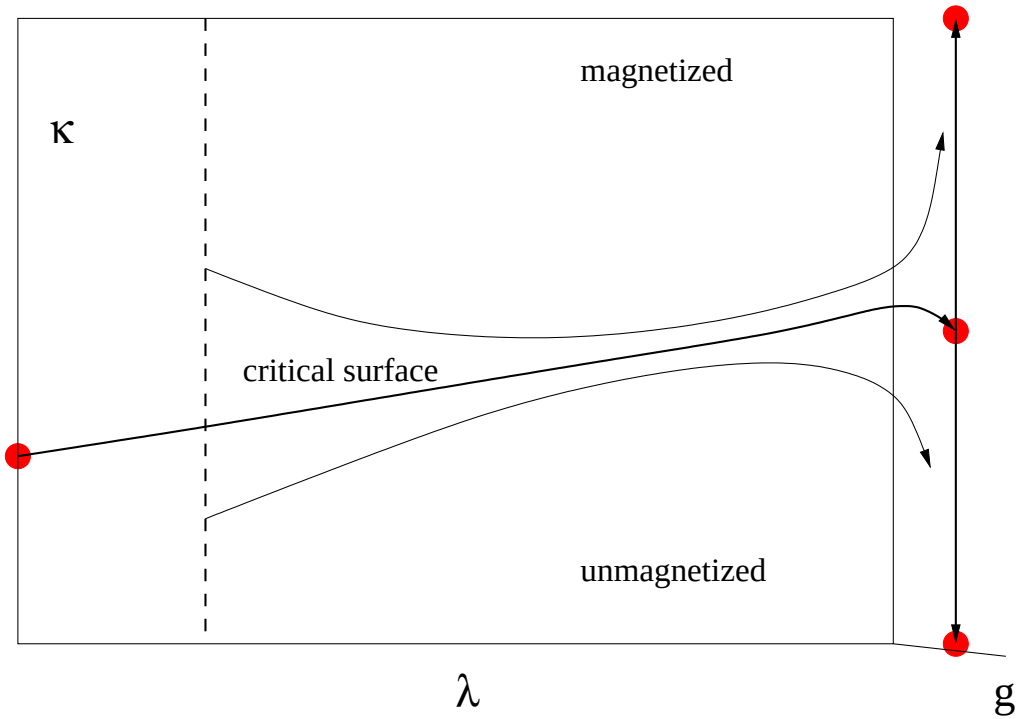
When the **hopping parameter** κ is large we may expand around the free field limit. This is lattice perturbation theory. In the limit when $\kappa \to 0$ we may make a hopping parameter expansion around a solution in which the sites are decoupled.
When $\lambda \to \infty$ the field at the scale of the cutoff must sit at the minimum of the potential, so the model looks like the Ising model. From our earlier discussion, we expect that the critical exponents of scalar field theory must be the same as that of the Ising model, *i.e.*, the two are in the same **universality class**.

Monte Carlo simulations

In general the theory is investigated by Monte Carlo simulations. The algorithm is the following—

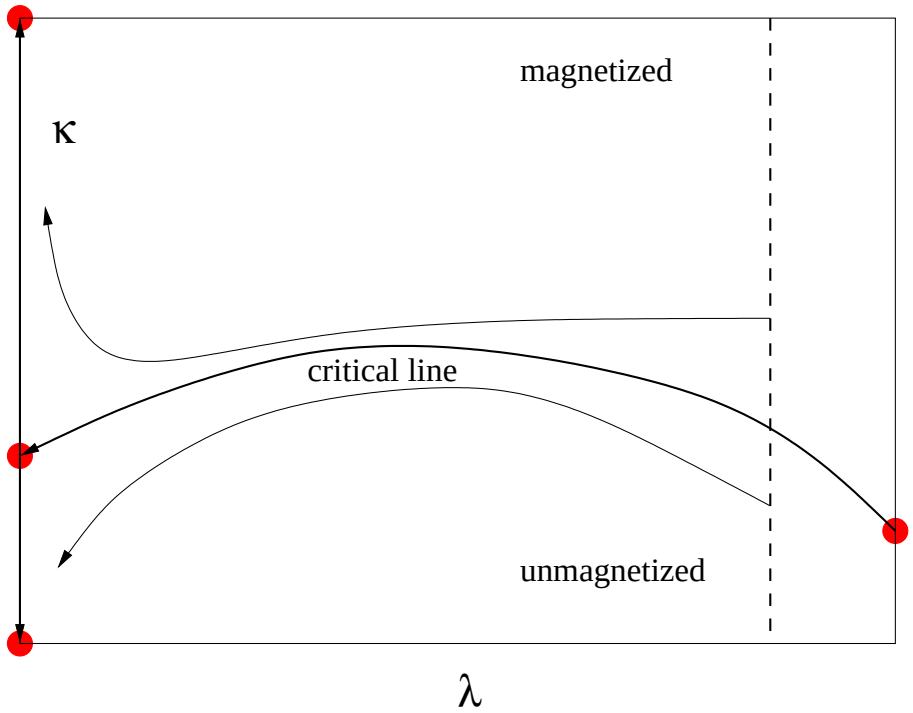
1. Start from a randomly generated **configuration of fields**, $\phi(x)$, on the lattice.
2. At one lattice site, x, make a random suggestion for a new value of the field, $\phi'(x)$.
3. Make a **Metropolis choice** as follows. If the change in the action, ΔS, due to the change in the field is negative, or $\exp(-\beta\Delta S)$ is smaller than a random number r (uniformly distributed between 0 and 1) then accept the suggestion. Otherwise reject it.
4. **Sweep** through every site of the lattice repeating steps 2, 3.
5. At the end of each sweep make **measurements** of the moments of the field variables.
6. Repeat from step 2 as many times as the computational budget allows.

Bosons in $D = 2$



The theory of interacting bosons in $D = 2$ has a non-trivial critical point corresponding to the Ising model. RG trajectories lying anywhere on the critical surface are attracted to this. Since the scalar field in $D = 2$ is dimensionless, an infinite number of couplings, g, in addition to κ and λ, need to be tuned to get to it.

Triviality of the Higgs in $D = 4$



In $D = 4$ the only attractive point on the critical surface has $\lambda = 0$. Since all RG trajectories are attracted to $\lambda = 0$, close to the continuum limit perturbation theory can be used to examine the beta-function. (M. Luscher and P. Weisz, *Nucl. Phys.*, B 290, 25, 1987).

References

- R. P. Feynman and Hibbs, *The Path Integral and its Applications*, McGraw Hill.
- R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press.
- D. J. Amit and V. Martin-Mayor, *Field Theory, the Renormalization Group, and Critical Phenomena*, World Scientific.
- G. Toulouse and P. Pfeuty, *Introduction to the Renormalization Group and its Applications*, Grenoble University Press.
- N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Interscience.

Outline

Notation for lattice theories

1. The lattice spacing will always be written as a except when we use units where $a = 1$. When quantities with different dimensions are equated, it will mean that each quantity is made dimensionless by multiplying by appropriate factors of a and a is set to 1. For example, $2\pi + m = 2\pi/a + m$.
2. In contexts where there is no confusion, we will use the notation x, y, *etc.*, to denote either a point in continuum space-time, or on the lattice.
3. If we need more clarity, then sites on the lattice will be denoted by vectors of integers $\mathbf{i}$, $\mathbf{j}$, *etc.*.
4. To specify links on the lattice, we need to specify the lattice point x and a direction μ. Directions will be denoted by Greek symbols μ, ν, *etc.*. Unit vectors in these directions will be written as $\hat{\mu}$, $\hat{\nu}$, *etc.*. As a result, the nearest neighbours of x are $x + \hat{\mu}a$.

Notation for groups

1. Group indices are denoted by Greek symbols α, β, *etc.*.
2. Generators of group algebra are denoted λ_α, and can be represented by traceless Hermitean matrices. They satisfy the algebra $[\lambda_\alpha, \lambda_\beta] = if_{\alpha\beta\gamma}\lambda_\gamma$, where f is called the structure constant. An algebra element, $A = A^\alpha \lambda_\alpha$, is a linear combination of the generators with A^α real.
3. Generators are normalized so that $\mathrm{Tr}\,(\lambda_\alpha \lambda_\beta) = \delta_{\alpha\beta}/2$
4. Members of the group are exponentials $U = \exp(iA)$. In the representation where A are traceless Hermitean, U are unitary, *i.e.*, $UU^\dagger = 1$. For example, the Wigner $\mathcal{D}$ matrices are representations of the rotation group $O(3)$.
5. Characters of group elements are $\chi_r(U) = \mathrm{Tr}\, U$, so they depend on the representation, r. For example, for rotation through a fixed angle, the traces of Wigner $\mathcal{D}$ matrices depend on the angular momentum L.

The center of a group

1. We define the **center of a group** to be those elements, z, which commute with all elements of the group, *i.e.*, $zU = Uz$ for any U.
2. The center of a group is non-empty because the identity is always a member of the center.
3. Since $zU = Uz$, we find that $U^{-1}z^{-1} = z^{-1}U^{-1}$ for any U in the group. So, if z is in the center then z^{-1} is also in the center.
4. The center elements obey all other properties which a group must have since they are elements of a larger group.
5. Hence the center of a group is an Abelian subgroup.

What is a gauge field?

Minimal coupling of a gauge field $A_\mu = A_\mu^\alpha \lambda_\alpha$ (λ_α are generators of the gauge group) means that the momentum operator is

$$(p_\mu - eA_\mu)\psi = (i\partial_\mu - eA_\mu)\psi.$$

So A_μ is involved in **parallel transporting** the wave-function by an infinitesimal amount in the direction μ. A change in $A_\mu(x) \to A_\mu(x) + \partial_\mu f(x)$, *i.e.*, a **gauge transformation** is unphysical and can be absorbed into an equally unphysical phase of ψ.

On a lattice there are no infinitesimal displacements. The derivative operator is replaced by a finite difference. The parallel transporter must be replaced by the analogous finite quantity. This is the (group valued) link variable

$$U_\mu(x) = \exp(iaA_\mu).$$

Gauge transformations

On a lattice we must now promote the algebra valued local gauge function $f(x)$ to a local group-valued field $V(x) = \exp[if(x)]$. The derivative of the function in a transformation means that we must use the function at two points. An obvious generalization is

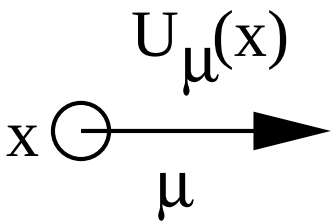
$$U_\mu(x) \to V(x)U_\mu(x)V^\dagger(x + \hat{\mu}a).$$

If we parallel transport a state across a path $(x_1, x_2, \cdots x_N)$, where $x_{l+1} = x_l + \hat{\mu}_l a$, then the relevant field and its gauge transform are—

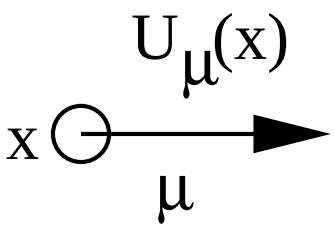
$$U(x_1, x_2, \cdots x_N) = \prod_{l=1}^{N-1} U_{\mu_l}(x_l) \to V(x_1)U(x_1, x_2, \cdots x_N))V^\dagger(x_N).$$

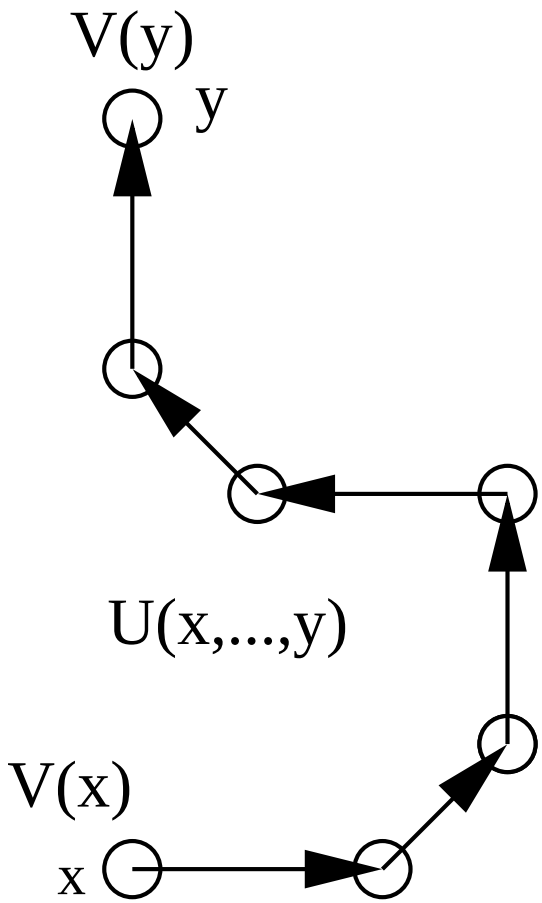
If we go around a closed loop ($x_1 = x_N$) then the gauge transformation is purely local. The trace of a closed loop $U(x_1, x_2, x_3, \cdots, x_1)$ is then gauge invariant. These are the only gauge invariant quantities that one can build.

Some pictures

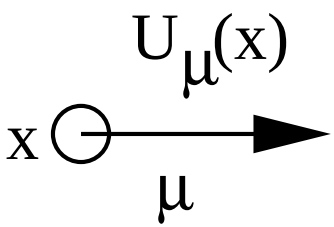


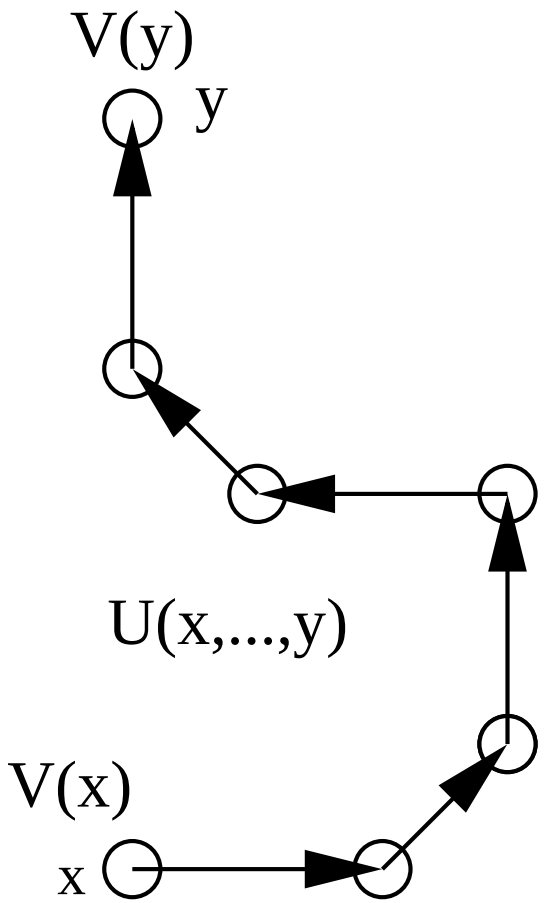
Some pictures

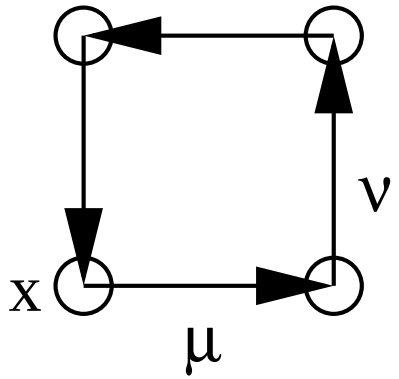




Some pictures







The Wilson action $S = \beta \sum_{x,\mu\leq\nu} \mathrm{Re}\,\mathrm{Tr}\,[P_{\mu,\nu}(x) - 1]$

The Wilson action is written in terms of a **plaquette**, which is the smallest loop on a hypercubic lattice—

$$\begin{aligned} P_{\mu\nu}(x) &= U(x, x+\hat{\mu}a, x+\hat{\mu}a+\hat{\nu}a, x+\hat{\nu}a, x) \\ &= U_\mu(x) U_\nu(x+\hat{\mu}a) U_\mu^\dagger(x+\hat{\nu}a) U_\nu^\dagger(x). \end{aligned}$$

To leading order in a the exponents of $U_\nu(x+\hat{\mu}a)$ and $U_\nu^\dagger(x)$ give the $\partial_\mu A_\nu$ term. Using the **Baker Campbell Hausdorff** formula

$$\mathrm{e}^x \mathrm{e}^y = \mathrm{e}^{x+y+[x,y]/2},$$

we recover the field commutators for non-Abelian fields. Putting all of it together

$$P_{\mu\nu}(x) = \exp\left[ia^2\partial_\mu A_\nu(x) - ia^2\partial_\nu A_\mu(x) + a^2[A_\mu(x), A_\nu(x)] + \mathcal{O}(a^4)\right].$$

The trace gives $1 + a^4 F_{\mu\nu}F^{\mu\nu} + \mathcal{O}(a^6)$, thus reproducing the continuum Yang-Mills' action.

The partition function $Z(\beta) = \int \prod_{x,\mu} dU_\mu(x) \exp[-S]$

The integrals in the partition function are Haar integrals over the gauge group. They are normalized and translation invariant so that

$$\int dU = 1, \qquad \int dU U = 0, \int dU f(U) = \int dU f(UV),$$

where V is a fixed group element. Under the Haar measure group characters are orthogonal, *i.e.*,

$$\int dU \chi_p^*(U) \chi_q(U) = \delta_{pq},$$

where p and q label representations, and the star denotes complex conjugation. Given a complex valued function on the group, $f(U)$, we can perform **harmonic analysis** (Fourier transforms) on the group

$$f_r = \int dU \chi_r^*(U) f(U).$$

A strong coupling expansion

In the limit $\beta \to 0$ one has $\exp(-S) \to 1$ so that $Z = 1$. The **strong coupling expansion** consists of corrections around this limit. This uses the following facts about group integrals—

$$\int dU \chi_r(U) = \delta_{r0}, \int dU \chi_r(VU)\chi_s(U^\dagger W) = \delta_{rs}\frac{1}{d_r}\chi_r(VW),$$

where d_r is the dimension of the representation r, and $r = 0$ is the trivial representation, $U = 1$.
For a plaquette, P, we use the observation that

$$\mathrm{e}^{-\beta P} = c_0(\beta)\left[1 + \sum_{r\neq 0} d_r a_r(\beta)\chi_r(P)\right],$$

where the functions a_r can be computed, and the leading power of β increases with r. We see that the only contributions to Z come from closed surfaces of plaquettes.

Other methods

Other methods available for dealing with the Wilson action are

1. An expansion around $\beta \to \infty$, *i.e.*, the weak coupling expansion. Such a **perturbation expansion** on the lattice has more vertices than the continuum expansion. As a result, high order computations become difficult. The most important result from lattice perturbation theory is the computation of the beta-function. This gives a check on the universality of the function up to two-loop order, and thereby a relation between $\Lambda_{\rm lat}$ and $\Lambda_{\overline{MS}}$.
2. The preferred methods are **numerical simulations**, either the Metropolis, heat-bath or over-relaxation methods. In a numerical computation the lattice must be finite, say N^D, so that there are both infrared and ultraviolet cutoffs. As a decreases, if the volume is unchanged in physical units, then N must increase. So the number of degrees of freedom increases as N^D, leading to a rapid increase in computer time

Does QCD confine?

- You do not build new formulations in order to answer the same old questions. The unanswered big question for perturbation theory is whether non-Abelian gauge theories confine. Wilson framed the question in terms of the potential between two static quarks, $V(r)$.
- In the 60's it was discovered that the known Hadron spectrum showed **Regge behaviour**, *i.e.*, $M^2 \simeq J$. This was shown to be the spectrum that arises from a spinning string of finite tension (although the string theory of hadrons was, and remains, inconsistent).
- In the 70's it was discovered that non-Abelian fields might not spread out from a charge in Coulomb's radial pattern, but might collapse into a **flux tube**. In that case one might have $V(r) \propto r$, as for a string rather than the $V(r) \propto 1/r$ assumed in perturbation theory.

Static quark sources

A quark source couples to gauge field through the term

$$\delta S = \int d^D x j^\mu A_\mu \to \int d^D x \delta^{(D-1)}(x) A_0 = \int dx_0 A_0,$$

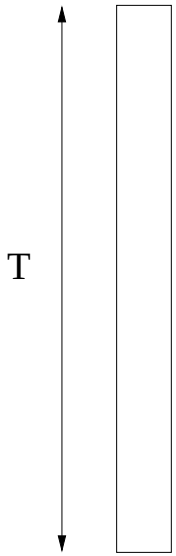
where the last result is obtained after taking the limit of a static quark sitting at the spatial point $x = 0$ which has only the component j_0. After a Wick rotation to Euclidean space, this part of the action reduces to a link in the time direction. The symmetry of Euclidean rotations can then be used to rotate these into any direction one chooses.

As a result, a static quark (or antiquark) source can be represented in the lattice theory as a sequence of gauge fields U along the path taken by the quark.

Wilson loops

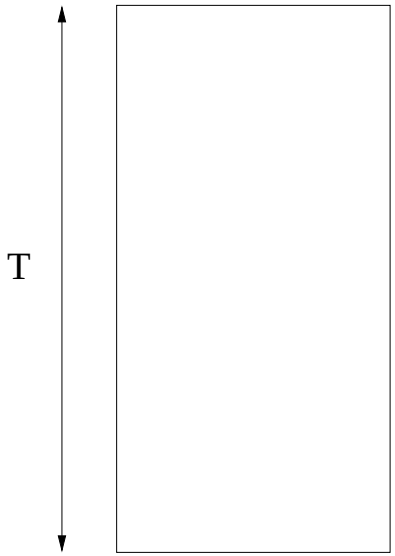
$$\begin{aligned}
\exp[-V(r)] &= \lim_{T\to\infty} \left\langle \mathrm{Tr}\, U(x, x+\hat{z}, \cdots, x+\hat{z}r, \cdots, \right. \\
&\qquad \left. x+\hat{z}r+\hat{t}T, \cdots, x+\hat{t}T, \cdots, x)\right\rangle \\
&= \lim_{T\to\infty} W(r,T).
\end{aligned}$$

Wilson loops



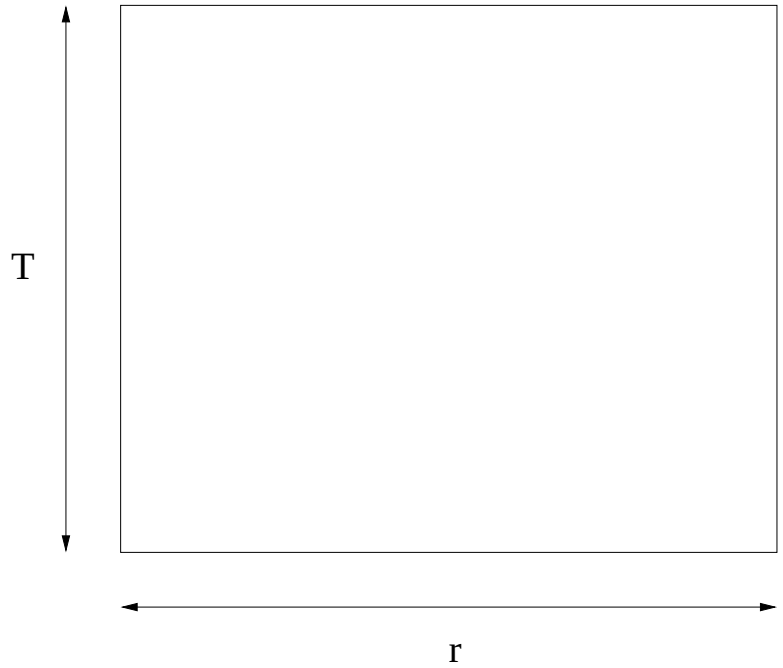
$$\begin{aligned}\exp[-V(r)] &= \lim_{T\to\infty} \left\langle \mathrm{Tr}\, U(x, x+\hat{z}, \cdots, x+\hat{z}r, \cdots, \right. \\ &\qquad \left. x+\hat{z}r+\hat{t}T, \cdots, x+\hat{t}T, \cdots, x)\right\rangle \\ &= \lim_{T\to\infty} W(r,T).\end{aligned}$$

Wilson loops



$$\begin{aligned}\exp[-V(r)] &= \lim_{T\to\infty} \Big\langle \mathrm{Tr}\, U(x, x+\hat{z}, \cdots, x+\hat{z}r, \cdots, \\ &\qquad x+\hat{z}r+\hat{t}T, \cdots, x+\hat{t}T, \cdots, x)\Big\rangle \\ &= \lim_{T\to\infty} W(r,T).\end{aligned}$$

Wilson loops



$$\begin{aligned}\exp[-V(r)] &= \lim_{T\to\infty} \Big\langle \mathrm{Tr}\, U(x, x+\hat{z}, \cdots, x+\hat{z}r, \cdots, \\ &\qquad x+\hat{z}r+\hat{t}T, \cdots, x+\hat{t}T, \cdots, x)\Big\rangle \\ &= \lim_{T\to\infty} W(r,T).\end{aligned}$$

Confinement and glueballs

If $V(r) \propto r$ then $\log W(r,T) \propto rT$. This is called the **area law**. This is obtained at strong coupling because the expectation value of a Wilson loop is proportional to the number of plaquettes it contains. If each plaquette expectation value is p, then $W(r,T) \simeq p^{rT}$.

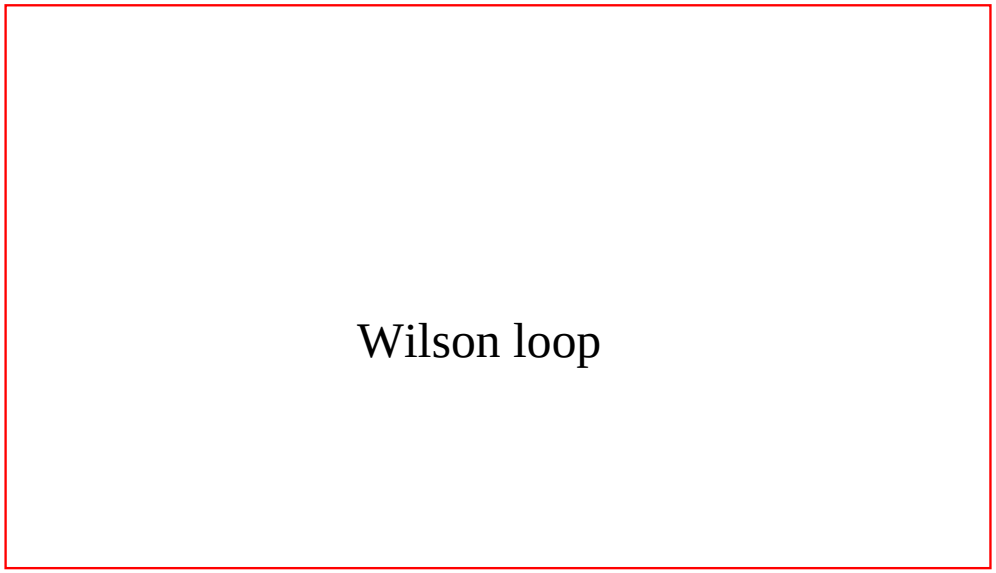
The correlation function between colour singlet operators is mediated by objects called glueballs. If $-\log W(r,T) \simeq \sigma rT$, *i.e.*, the string tension is σ, then the correlation function of L^2 sized loops separated by distance r is given by

$$C(r,L) \simeq \exp(-4\sigma rL),$$

i.e., the glueball correlations fall exponentially, and the glueball mass is $4\sigma L$.

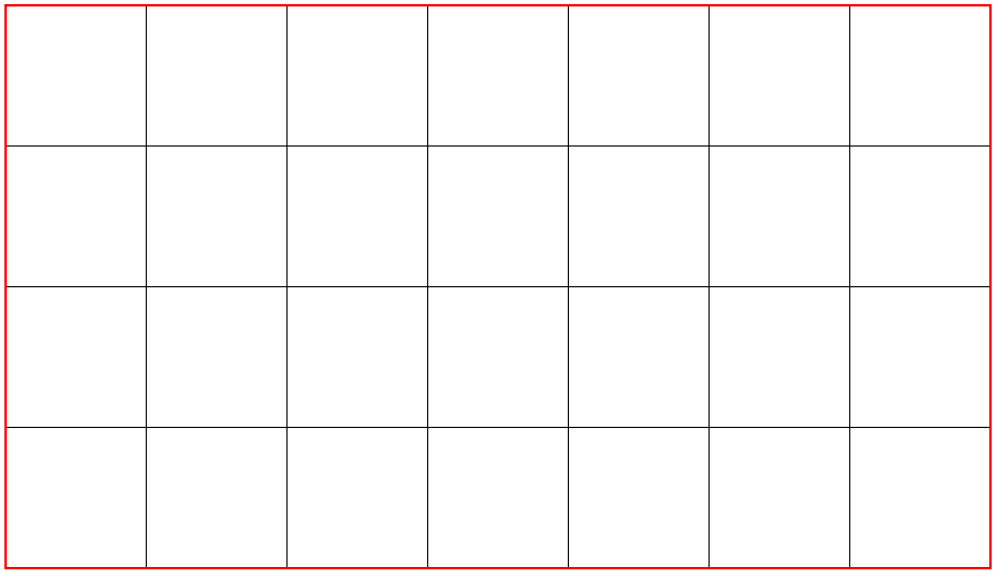
This is not a proof of confinement for QCD because the strong coupling phase does not have a continuum limit due to **string roughening**.

The strong coupling argument



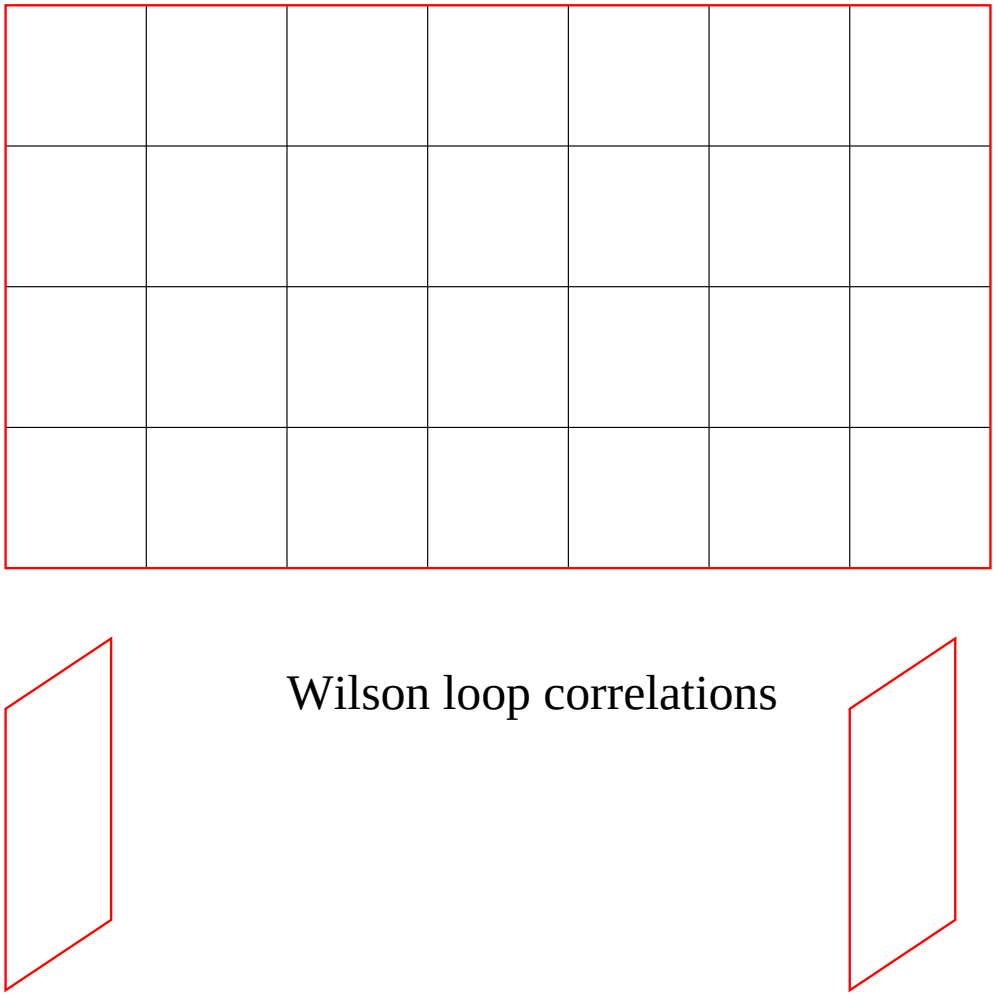
J.-M. Drouffe and C. Itzykson, *Phys. Rep.*, 38, 133, 1978

The strong coupling argument



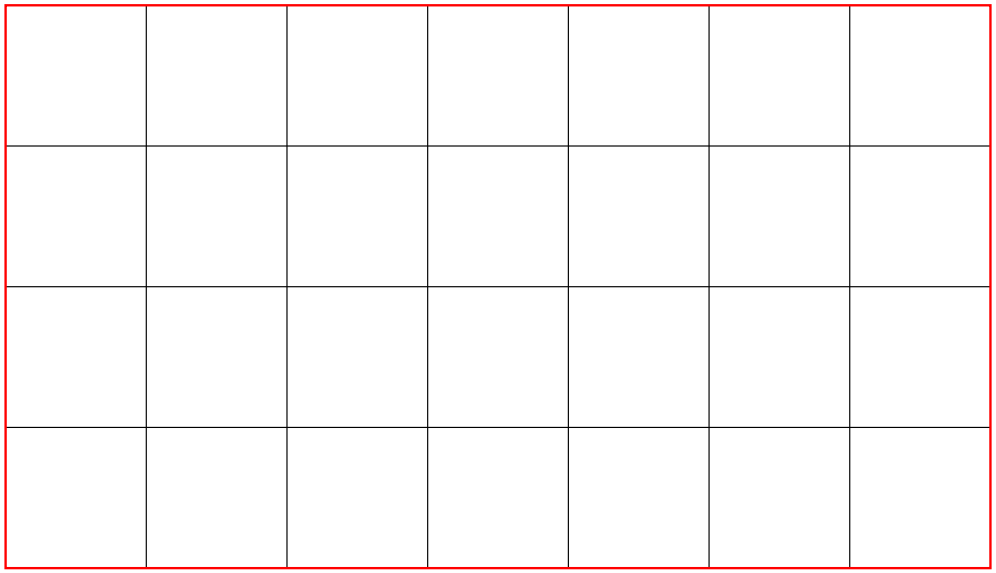
J.-M. Drouffe and C. Itzykson, *Phys. Rep.*, 38, 133, 1978

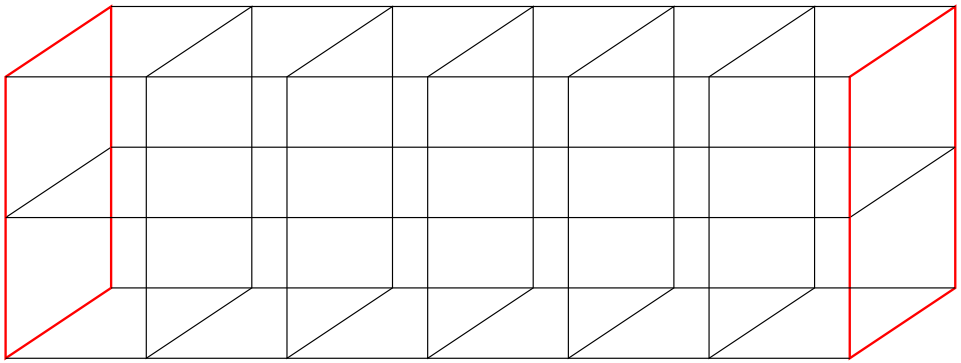
The strong coupling argument



J.-M. Drouffe and C. Itzykson, *Phys. Rep.*, 38, 133, 1978

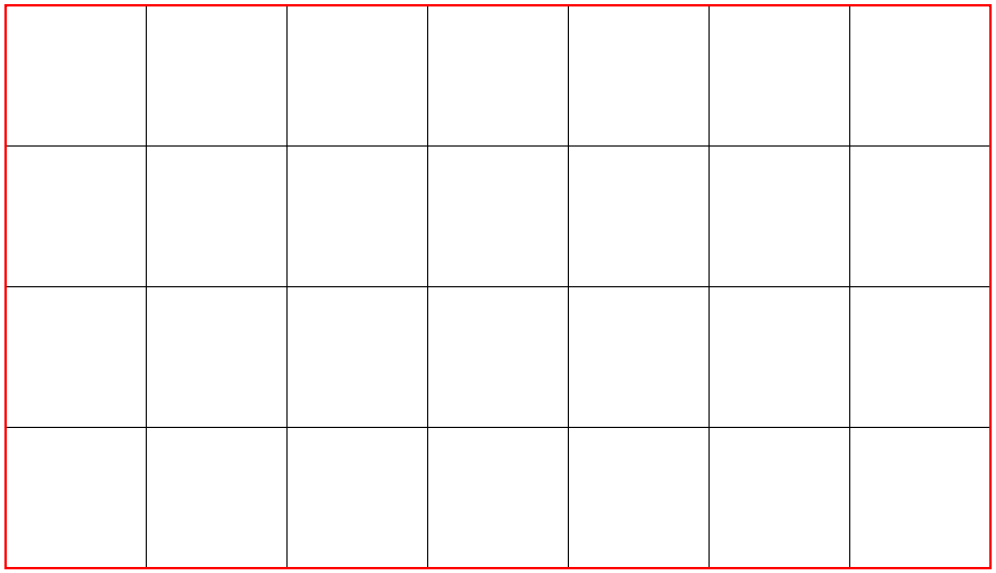
The strong coupling argument

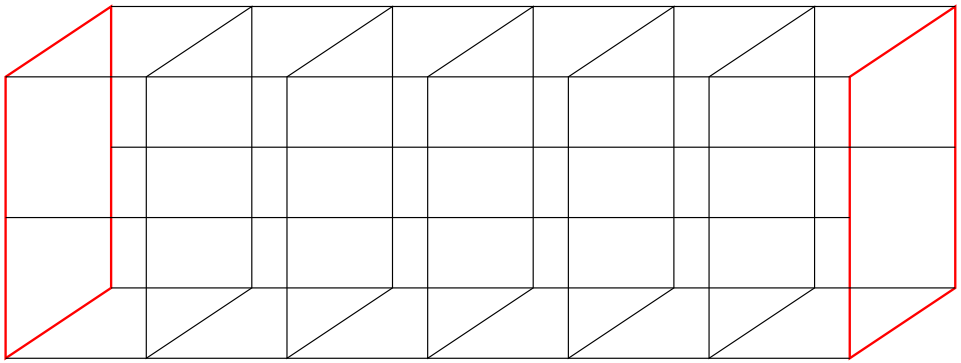




J.-M. Drouffe and C. Itzykson, *Phys. Rep.*, 38, 133, 1978

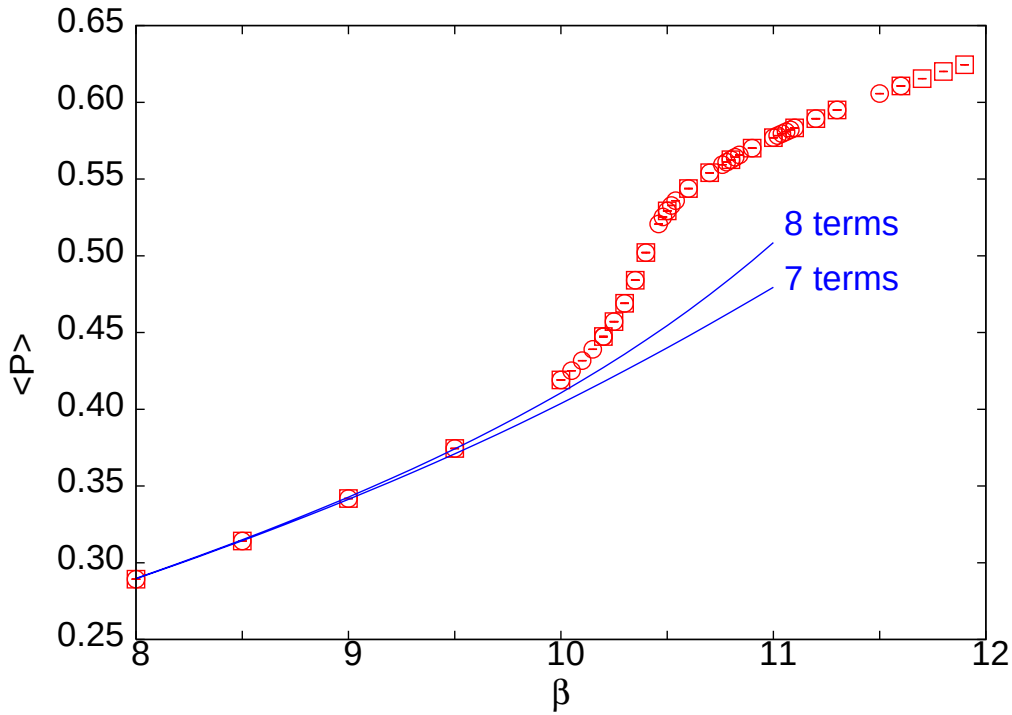
The strong coupling argument





J.-M. Drouffe and C. Itzykson, *Phys. Rep.*, 38, 133, 1978

Breakdown of the strong coupling expansions



S. Datta and S. Gupta, *Phys. Rev.*, D80, 114504, 2009

Comparison of the strong coupling expansion for the plaquette and Monte Carlo measurements for SU(4) gauge theory on a 16^4 lattice. The strong coupling expansion seems to break down at $\beta \simeq 10$. Fluctuations of the surface become unbounded.

Gauge theories at finite temperature

The statistical mechanics of a gauge theory is examined by evaluating the partition function with periodic boundary conditions in the Euclidean time direction and sending the spatial size to infinity (the **thermodynamic limit**). In practice one computes on a $N_t \times N_s^{D-1}$ lattice with $T = 1/(aN_t)$ and limit $L = N_s a \gg 1/T$. At finite temperature the action has a global symmetry under multiplication of time-like links by any element of the center, of the gauge group *i.e.*,

$$\text{if } U'_t(x) = zU_t(x), \quad \text{then } S[U'] = S[U].$$

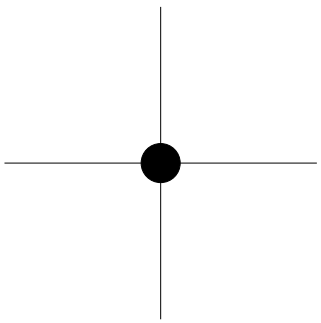
To check this, note that if $U \to zU$ then $U^\dagger \to z^{-1}U^\dagger$. As a result each plaquette is invariant, and hence the action is invariant.
If the center of the gauge group is non-trivial (*i.e.*, contains more than just the identity) then one can investigate whether this global symmetry is broken or restored as the temperature changes.

An order parameter for deconfinement

The free energy, F, of a single static quark is given by the **Polyakov loop**

$$L = \mathrm{Tr} \prod_{I} U_t(x + \hat{t}I), \quad \mathrm{e}^{-F} = \mathrm{Re}\,\langle L \rangle\,.$$

L does not go into itself under a center transformation and hence can serve as an order parameter for the breaking of the center symmetry. At $T = 0$ $\langle L \rangle = 0$, since a single static quark has infinite free energy in the confined vacuum. However, if there is deconfinement, then $\mathrm{Re}L$ may be non-zero. This is essentially what happens above a **transition temperature** for deconfinement, T_c.



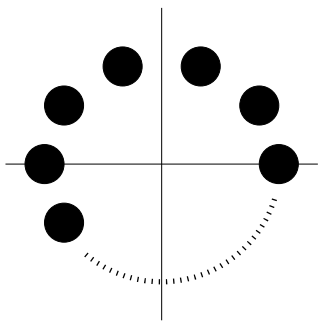
However, $\langle L \rangle = 0$ identically because of the center symmetry of the action. As a result, for $T > T_c$, L can take several non-zero values, all of which are center transforms of each other. The sum over these vanishes.

An order parameter for deconfinement

The free energy, F, of a single static quark is given by the **Polyakov loop**

$$L = \mathrm{Tr} \prod_{I} U_t(x + \hat{t}I), \quad \mathrm{e}^{-F} = \mathrm{Re}\,\langle L \rangle\,.$$

L does not go into itself under a center transformation and hence can serve as an order parameter for the breaking of the center symmetry. At $T = 0$ $\langle L \rangle = 0$, since a single static quark has infinite free energy in the confined vacuum. However, if there is deconfinement, then $\mathrm{Re}L$ may be non-zero. This is essentially what happens above a **transition temperature** for deconfinement, T_c.



However, $\langle L \rangle = 0$ identically because of the center symmetry of the action. As a result, for $T > T_c$, L can take several non-zero values, all of which are center transforms of each other. The sum over these vanishes.

The equation of state

The classic method for computing the equation of state is to obtain operators whose expectation values give thermodynamic variables. The definitions are

$$E = \frac{T^2}{V} \left. \frac{\partial \log Z}{\partial T} \right|_V, \; P = T \left. \frac{\partial \log Z}{\partial V} \right|_T .$$

These derivatives give combinations of plaquettes. Such operators seem to have strong finite lattice artifacts.
As a result, this operator method is now superceded by the integral method. In this new method, one uses the fact that

$$P = \frac{T}{V} \log Z = P_0 + \frac{T}{V} \int_{\beta_0}^{\beta} d\beta \, T^4 \frac{\partial \log Z}{\partial \beta}.$$

In addition, one uses the operator expression for $E - 3P$ to obtain the complete thermodynamics. (G. Boyd *et al.*, *Nucl. Phys.*, B469, 419, 1996).

The Metropolis algorithm and acceptance rates

The Metropolis and Heat-Bath algorithms continually bring each degree of freedom into equilibrium with its neighbours at a given β. For every link with value U, one suggests a random new value U' and accepts it with the Metropolis probability

$$p = \min\left[1, \mathrm{e}^{\beta \Delta S}\right], \quad \text{where } \Delta S = S[U'] - S[U].$$

Some measure of the distance between U and U', $\delta = |U' - U'|$ is tuned so that the acceptance rate in equilibrium, $\langle p \rangle_\beta$, is around 80%. Extremely high or low acceptance rates mean that the movement in configuration space is very slow. It would be an useful exercise to measure $\langle p \rangle_\beta$ as a function of δ and try to find invariance principles as the bare coupling, β is changed.

The heat-bath algorithm

If all neighbours of a single link are kept fixed, while it is repeatedly updated by Metropolis, then asymptotically it reaches a certain "thermal" distribution. Heat-bath (HB) algorithms are set up to sample this in one step.
For an Ising model if the spin of interest is s, and the sum over all neighbouring spins is h, then the limiting probabilities are

$$p(s) = \frac{\mathrm{e}^{-\beta h s}}{2\cosh \beta h}.$$

The Ising HB is simply to set $s = \pm 1$ with the above probabilities. For the U(1) group there is an HB due to Bunk (B. Bunk, unpublished), a fast HB specific to SU(2) (A. D. Kennedy and B. J. Pendleton, *Phys. Lett.*, B156, 393, 1985), and a method for extending this to any SU(N) (N. Cabibbo and E. Marinari, *Phys. Lett.*, B119, 387, 1982).

Critical slowing down

The rate at which the whole lattice moves through configuration space can be observed by a **damage spreading** exercise. (S. Gupta, *Nucl. Phys.*, B370, 741, 1992)

Make a copy of a configuration, C, and change it by a large amount in exactly one link. Call this configuration C'. Now subject C and C' to Monte Carlo evolution and compare them at the end of each sweep. The damage front is that region of the lattice through which the difference between the two configurations is much larger than δ. For the Metropolis and Heat-Bath algorithms this spread is diffusive, *i.e.*, the radius of the damage front changes as $\sqrt{t}$. This means that measurements are strongly correlated over time scales proportional to R^2.

The damage front stalls when $R \simeq \xi$, so global (thermodynamic) averages are correlated at time scales of the order of ξ^2.

The Over-Relaxation Algorithm

In the over-relaxation (OR) algorithm we move the link as much as possible without changing the action, *i.e.*, we solve $S[U'] = S[U]$ for the new value, U'. A new value is accepted with the probability dU'/dU, *i.e.*, in the ratio of the Haar measures at the two points. This ensures **detailed balance**. (OR)
Since OR does not change the action, one has to intersperse it with an occasional heat-bath (HB) sweep. Typically one takes N_{or} sweeps of OR per sweep of HB. A damage spreading computation shows that the damage radius grows linearly with the number of sweeps as long as N_{or} is of the order of the correlation length. (U. Wolff, *Phys. Lett.*, B288, 166, 1992)

Renormalized couplings

The bare coupling of SU(N) gauge theory is given by $g_B^2 = 2N/\beta$. The renormalized coupling, g, can be found by measurement of some reference quantity on the lattice and using its perturbation expansion to define g. This reference is often taken to be the plaquette since it is easy to measure. (G. P. Lepage, P. B. Mackenzie, *Phys. Rev.*, D48, 2250, 1993)

Such a measurement gives g at the scale of a. As one changes β one changes a and hence the renormalized coupling g^2. The beta-function measures the change in g^2 with changing a. When a is small enough, the renormalized coupling should be so small that one can use 2-loop beta-function

$$a\frac{dg}{da} = -\beta_0 g^3 - \beta_1 g^5.$$

A strategy for testing the continuum limit

Integrating the 2-loop beta-function one has

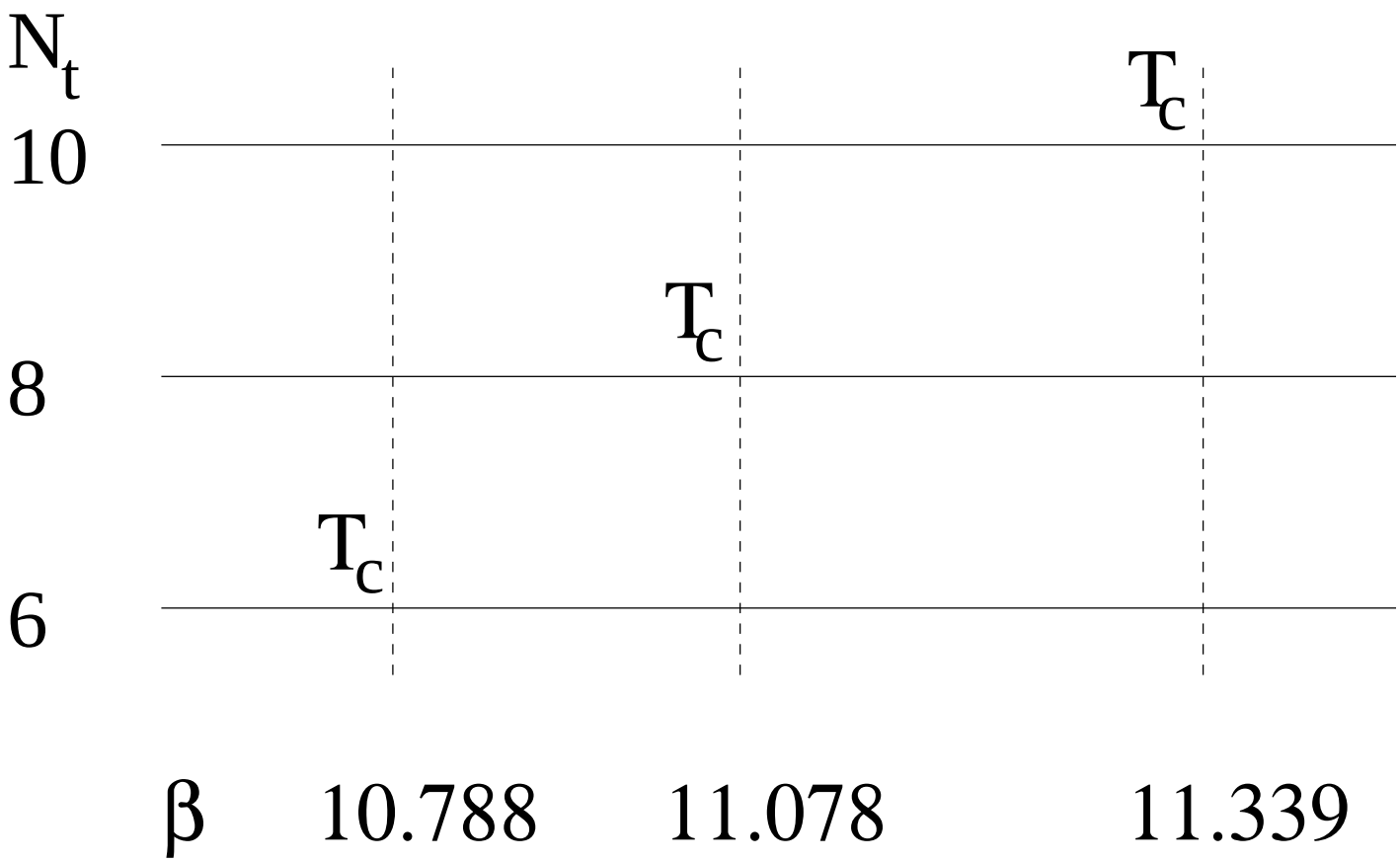
$$a\Lambda = kR\left(\frac{1}{4\pi\beta_0\alpha_S}\right),\ R(x) = \mathrm{e}^{-x/2}x^{\beta_1/(2\beta_0)}.$$

Find the bare coupling, β_c, at which a lattice with fixed N_t shows the deconfinement transition. At this bare coupling $a = 1/(N_t T_c)$. Measure α_S at this bare coupling. Then change N_t to N_t', measure the new critical bare coupling, β_c' and corresponding renormalized coupling α_S'. The ratio of lattice spacings then gives

$$\frac{N_t'}{N_t} = \frac{R[1/(4\pi\beta_0\alpha_S)]}{R[1/(4\pi\beta_0\alpha_S')]}.$$

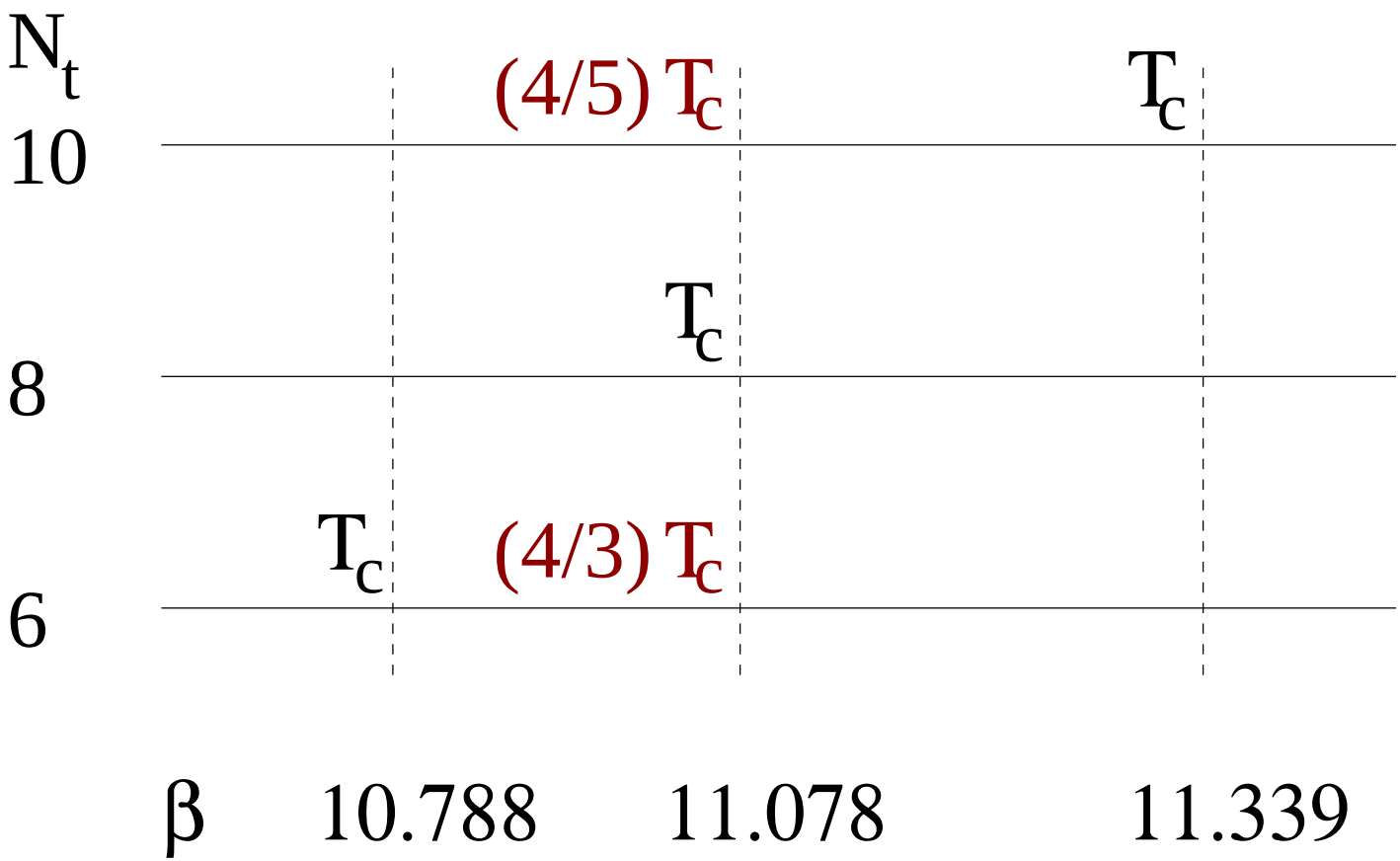
By increasing N_t in steps, one can find how small the lattice spacing has to be before 2-loop scaling begins to work.

Reaching the continuum limit



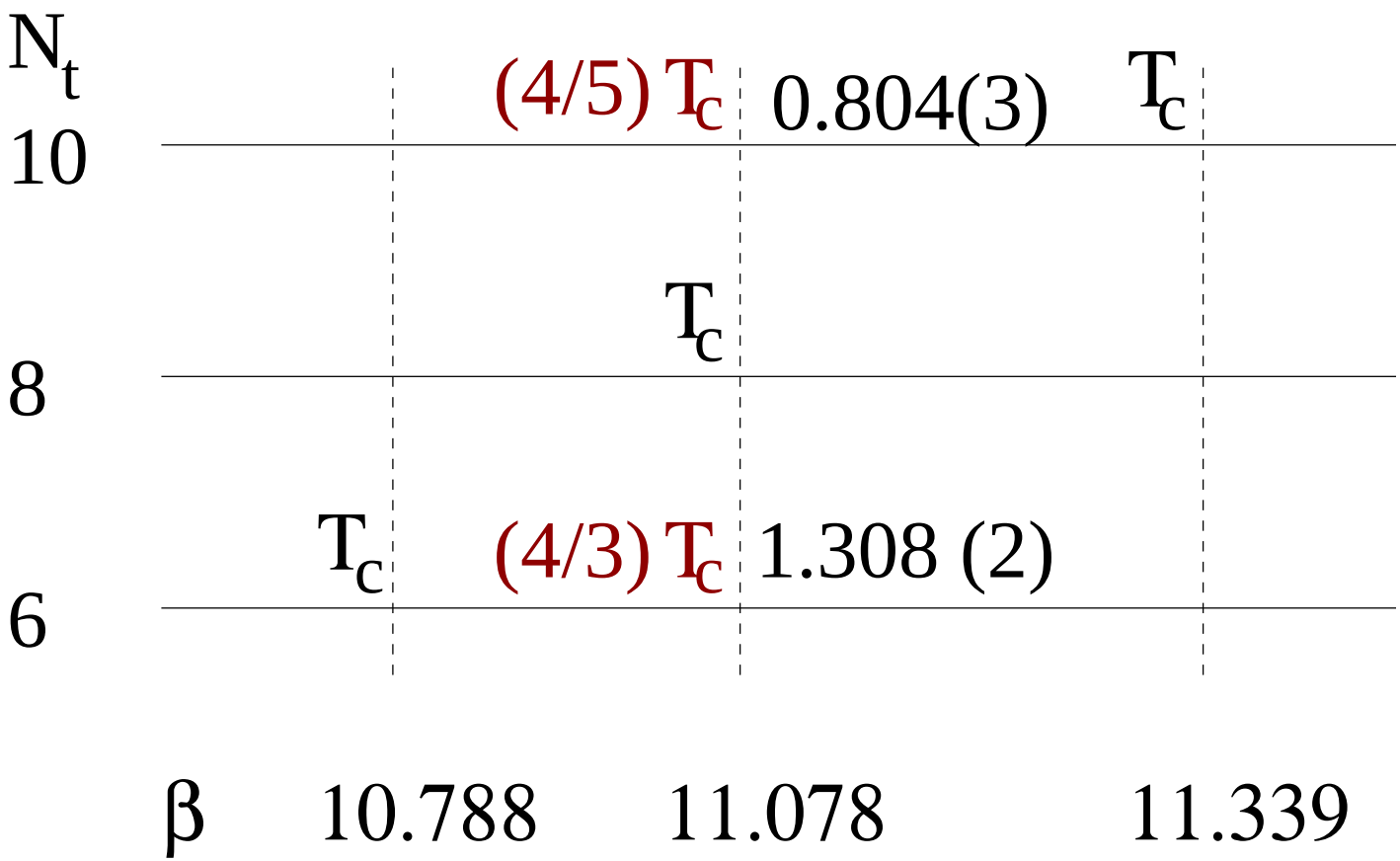
β_c determined to one part in 10^4; α_S from plaquette. 2-loop RG works with precision of two parts in 10^3 for $a \leq 1/(8T_c)$. (S. Datta and S. Gupta, *Phys. Rev.*, D80, 114504, 2009).

Reaching the continuum limit



β_c determined to one part in 10^4; α_S from plaquette. 2-loop RG works with precision of two parts in 10^3 for $a \leq 1/(8T_c)$. (S. Datta and S. Gupta, *Phys. Rev.*, D80, 114504, 2009).

Reaching the continuum limit



β_c determined to one part in 10^4; α_S from plaquette. 2-loop RG works with precision of two parts in 10^3 for $a \leq 1/(8T_c)$. (S. Datta and S. Gupta, *Phys. Rev.*, D80, 114504, 2009).

Outline

Euclidean Dirac Fermions in $D = 4$

The Dirac Hamiltonian in Euclidean $D = 4$ space-time, acting on 4-component Dirac spinors, $\Psi(x)$, is

$$H = m\beta - i\alpha_j \partial_j, \text{ where } \{\beta, \alpha_j\} = 0, \ \{\alpha_i, \alpha_j\} = 2\delta_{ij},$$

the braces denote anti-commutators, and Latin subscripts run over the three spatial directions. We choose β and α_j to be Hermitean. On the lattice we replace the derivative operator by the forward difference as before, and find

$$H = \sum_x m\Psi^\dagger(x)\beta\Psi(x) + \frac{i}{2}\sum_j \left[\Psi^\dagger(x+\hat{j})\alpha_j\Psi(x) - \Psi^\dagger(x)\alpha_j\Psi(x+\hat{j})\right].$$

Fourier transforms block diagonalize the Hamiltonian and give

$$H = \frac{1}{V}\sum_k \psi^\dagger(k) M \psi(k), \text{ where } M = m\beta + \sum_j \alpha_j \sin k_j.$$

Fermion doubling

These 4×4 Dirac blocks can be diagonalized by noting that $M^\dagger = M$ and that $MM^\dagger$ is diagonal. The eigenvalues are

$$E_k = \pm \sqrt{m^2 + \sum_j \sin^2 k_j}.$$

In the limit $k \to 0$ one finds the correct dispersion relation $E_k = \pm\sqrt{m^2 + k^2}$. However, $\sin^2 k_j$ vanishes not only when $k_j = 0$ but also at each edge of the Brillouin zone, *i.e.*, $k_j = \pm\pi$. Therefore, each corner of the Brillouin zone contains one copy of the Dirac Fermion with the correct dispersion relation. This is **Fermion doubling**.
Keeping the quadratic form $H = \sum_{xy} \Psi^\dagger(y) M \Psi(x)$, the degeneracy can be lifted by changing the **Dirac operator** M.

Wilson fermions

Wilson suggested adding irrelevant terms to M, for example,

$$\delta M = 3r\beta\delta_{xy} - \frac{r}{2}\beta\sum_j(\delta_{x+\hat{j},y} + \delta_{y+\hat{j},x}),$$

with $0 < r \leq 1$. Diagonalizing again by using Fourier transforms and the properties of the Dirac matrices, one has

$$E_k^2 = \left[m + r\sum_j(1 - \cos k_j)\right]^2 + \sum_j \sin^2 k_j.$$

For r in the above range, the degeneracies at the corners of the Brillouin zone are lifted. The dispersion relation at small k is

$$E_k^2 = m^2 + (1 + 2mr)k^2 + \mathcal{O}(k^4),$$

so that the corrections are of order a rather than a^2. Symanzik improvement becomes an important issue.

Spin-diagonalization of fermions

The action, S, for fermion fields $\Psi(x)$ is given by $\sum_{xy} \overline{\Psi}(x) M(x,y) \Psi(y)$, where the Dirac operator M is

$$M(x,y) = m\delta_{xy} + \frac{1}{2}\sum_{\mu} \gamma_\mu (\delta_{x+\hat{\mu},y} - \delta_{x-\hat{\mu},y}),$$

and $\gamma_0 = \beta$ and $\gamma_k = -i\gamma_0\alpha_k$. We can make a change of variables called **spin diagonalization**, $\Psi(x) = A(x)\psi(x)$, where $A(x)$ are 4×4 unitary matrices such that

$$A^\dagger(x)\gamma_\mu A(x+\hat{\mu}) = \Delta_\mu(x),$$

The choice $A(x) = \gamma_0^{x_0}\gamma_1^{x_1}\gamma_2^{x_2}\gamma_3^{x_3}$, where the x_i are (integer) components of x, gives a representation of Dirac matrices which are multiples of identity—

$$\Delta_\mu(x) = \prod_{i>\mu} (-1)^{x_i}.$$

Staggered (Kogut-Susskind) fermions

The transformed Dirac operator becomes

$$M(x,y) = m\delta_{xy} + \frac{1}{2}\sum_{\mu}\alpha_{\mu}(x)(\delta_{x+\hat{\mu},y} - \delta_{x-\hat{\mu},y}),$$

which is a multiple of the identity. Hence one can thin the degrees of freedom and keep only a single component of the field at each site. The 16 1-component fermions at the corners of the Brillouin zone can be interpreted as 4 **tastes** of 4-component fermions. The Dirac components have been distributed across 2^4 sites of the lattice which collapse to a single point in the continuum.
In the limit $m \to 0$ the field on the odd sublattice ($\sum_{\mu} x_{\mu} = \text{odd}$) connects only to that on the even lattice. Hence there is an exact global $U(1) \times U(1)$ chiral symmetry

$$\psi(x \in \text{odd}) \to U_o\psi(x),\ \overline{\psi}(x \in \text{odd}) \to \overline{\psi}(x)U_e^{\dagger},$$

and U_o and U_e interchanged on the even sublattice.

Spin-flavour decomposition

Introduce lattice coordinates $x_\mu = 2y_\mu + u_\mu$ where $u_\mu = 0$ or 1. Quark fields can be defined as

$$q_{\alpha a}(y) = \frac{1}{8}\sum_u \Gamma_{\alpha a}(u)\psi(2y+u), \qquad \Gamma(u) = \prod_\mu \gamma_\mu^{u_\mu}.$$

Here the index α refers to flavour space and a to spin (Dirac). Using the notation $\gamma \otimes \sigma$ for the direct product of a Dirac and flavour matrix ($t_\mu = \gamma_\mu^T$), ∇ for the forward derivative and δ for the second derivative, we find

$$\frac{M}{16} = m1 \otimes 1 + \sum_\mu (\gamma_\mu \otimes 1\nabla_\mu - \gamma_5 \otimes t_5\delta_\mu).$$

H. Kluberg-Stern, A. Morel, O. Napoli and B. Petersson, *Nucl. Phys.*, B220, 447, 1983

In the $a \to 0$ limit, the δ_μ term vanishes, and the chiral symmetry is enhanced to $U(4) \times U(4)$.

Some general considerations

The Fermion action is generally the quadratic form with a 4×4 matrix M,

$$S = \frac{1}{V} \sum_p \Psi(p) M(p) \Psi(p),$$

where the inverse of M is the fermion propagator. The zeroes of M correspond to the poles of the propagator, and give the particle content of the action. **Chiral symmetry** implies that in the massless case the transformation $\Psi \to \exp(i\alpha\gamma_5)\Psi$ would lead to $M \to -\gamma_5 M \gamma_5$. In addition one has hypercubic (rotational) symmetry and reflection positivity.
One makes the additional technical assumption of the locality of $M(x, y)$, which translates to a statement of $F(p)$ falling sufficiently fast at large p, so that the propagator is continuous. Then with periodic (or anti-periodic) boundary conditions, $M(p)$ must be periodic on the Brillouin zone and therefore have 16 zeroes.

No neutrinos on the lattice

The Nielsen-Ninomiya theorem states that if the lattice Hamiltonian for Weyl fermions satisfies

1. translation invariance,
2. locality (the Fourier transform of the kernel has continuous derivatives)
3. Hermiticity
4. any exactly conserved charges are local, have discrete quantum numbers and have bilinear currents

then there are equal number of left-handed and right-handed particles for every value of the charge.

Why inversion?

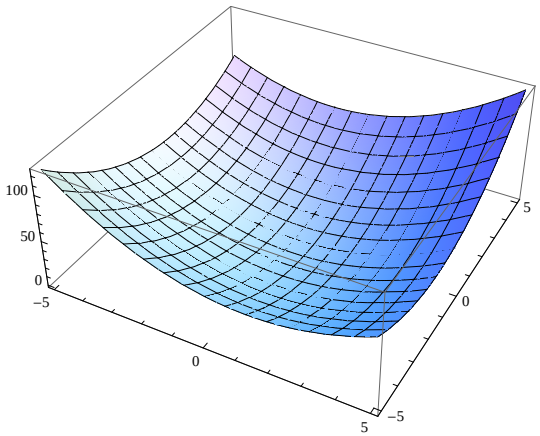
Meson correlation functions are objects like

$$C_\pi(x) = \left\langle \overline{\psi}(x)\gamma_5\psi(x)\overline{\psi}(0)\gamma_5\psi(0)\right\rangle = \left\langle \mathrm{Tr}\,\gamma_5 M^{-1}(0,x)\gamma_5 M^{-1}(x,0)\right\rangle .$$

Since one cannot program Grassman valued sources efficiently, one cannot write down local operators whose correlators will be saturated by meson states. Instead, every fermionic measurement requires finding the inverse of the Dirac operator.
There are fast matrix inversion methods which scale as the third power of the size of the matrix, N. However, since the size of the Dirac operator is proportional to the number of sites on the lattice, these generic methods are too slow for the fermion problem.
Instead one utilizes the fact that the Dirac operator is very sparse (since it is essentially a first derivative operator).
Sparse matrices can be dealt with very efficiently, for example inverting a tri-diagonal matrix is linear in N.

Changing the problem

The fastest way to solve $Ax = b$ is the **conjugate gradient** (CG) algorithm, if A is Hermitean and positive definite (M is not, but $M^\dagger M$ is, so we solve $M^\dagger Mx = M^\dagger b$ instead of $Mx = b$). What we actually do by the CG is to solve the equivalent problem of minimizing $f(x) = x^\dagger Ax/2 - b^\dagger x + c$. This approach follows Shewchuk's wonderful introduction [5].



The quadratic form plotted here is obtained with $b = (0, 0)$, $c = 0$ and

$$A = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}.$$

The solution of $Ax = 0$ is $x = (0, 0)$, which is where the minimum lies. The eigenvalues of A are 5 and 3, the eigenvector corresponding to the former is $v_5 = (1, 1)/\sqrt{2}$ and to the latter is $v_3 = (1, -1)/\sqrt{2}$.

The steepest descent method

We will construct an iterative process which will be stopped as soon as the solution, x^i, is good enough for use. The **error** is $\epsilon^i = x^* - x^i$ where x^* is the solution and the **residual** is $r^i = b - Ax^i$. Clearly, $r^i = A\epsilon^i$. Importantly, $r^i = -f'(x^i)$, so the residual is the direction of the **steepest descent**.
The **method of steepest descent** is simple. If one has reached x^i then, by the above argument, $x^{i+1} = x^i + \alpha r^i$. The value of α is chosen to reach the minimum along a line, *i.e.*, $f'(x^{i+1}) \cdot r^i = 0$.
For the quadratic form, one has

$$0 = r^i \cdot (b - Ax^{i+1}) = r^i \cdot (r^i - \alpha A r^i), \text{ so } \alpha = \frac{r^i \cdot r^i}{r^i \cdot Ar^i}.$$

Finally, we have $r^{i+1} = b - Ax^{i+1}$, and the recursion is completely set up. Matrix-vector multiplication (MvM) twice per iteration!
Using $p = Ar^i$ and $r^{i+1} = r^i - \alpha p$, MvM once per iteration.

Convergence

If the eigenvectors of A are v_k (assumed normalized) with eigenvalues λ_k, then we can write

$$\epsilon^i = \sum_k \xi_k v_k, \qquad r^i = A\epsilon^i = \sum_k \xi_k \lambda_k v_k.$$

Using this expansion in eigenvectors, we can write

$$\alpha = \frac{\sum_k \xi_k^2 \lambda_k^2}{\sum_k \xi_k^2 \lambda_k^3}.$$

The recursion relation for r^i then gives

$$\epsilon^{i+1} = \sum_k \xi_k (1 - \alpha\lambda_k) v_k, \qquad r^{i+1} = \sum_k \xi_k \lambda_k (1 - \alpha\lambda_k) v_k.$$

Defining $||v||^2 = v.Av$, we find

$$||\epsilon^{i+1}||^2 = ||\epsilon^i||^2 \omega^2 \text{ where } \omega^2 = 1 - \frac{(\sum_k \xi_k^2 \lambda_k^2)^2}{(\sum_k \xi_k^2 \lambda_k^3)(\sum_k \xi_k^2 \lambda_k)}.$$

The conjugate directions algorithm

In the steepest descent method our search direction was r^i. We switch to a set of mutually conjugate p^i, *i.e.*, $p^i \cdot A p^j = \delta_{ij}$. Taking the iteration $x^{i+1} = x^i + \alpha p^i$, the minimization condition becomes $r^{i+1} \cdot p^i = 0$. As a result, one finds

$$\alpha_i = \frac{p^i \cdot r^i}{p^i \cdot A p^i} = \frac{p^i \cdot A \epsilon^i}{p^i \cdot A p^i}.$$

Note that $\epsilon^{i+1} = \epsilon^i - \alpha p^i$. Does the iteration converge in N steps? Decompose $\epsilon^0 = \sum \xi_k d_k$. Since p^k are mutually conjugate, we find

$$\xi_k = \frac{p^k \cdot A \epsilon^0}{p^k \cdot A p^k} = \frac{p^k \cdot A(\epsilon^0 + \sum_{j<k} \alpha_j p^j)}{p^k \cdot A p^k} = -\alpha_k.$$

So, the conjugate directions method cuts away the components of ϵ^0 one by one, and converges in N steps.

The conjugate gradient construction

The CG corresponds to constructing the p^i from the set of r^i already generated. If so, and previous steps gave residuals r^0, r^1, $\cdots$, r^{i-1}, then r^i is orthogonal to the subspace spanned by them, since $p^i \cdot r^j = p^i \cdot Ae^j = 0$ if $i < j$. As a result, the r^i are orthogonal to each other, so r^j is always a new search direction. Since r^i are linear combinations of previous residuals and Ap^i, the subspace spanned by them is also spanned by r^0, Ar^0, A^2r^0, *etc.*. This is called a **Krylov space**.

The conjugate directions can be constructed by a Gram-Schmidt process (with "metric" A) in general, but because of the orthogonalities here, one step suffices. As a result

$$p^{i+1} = r^{i+1} + \beta_{i+1} p^i, \text{ where } \beta_{i+1} = \frac{r^{i+1} \cdot r^{i+1}}{r^i \cdot r^i}.$$

The immense simplification is that the previous vectors do not have to stored for Gram-Schmidt conjugation

The conjugate gradient algorithm

Putting everything together, the algorithm is initialized with $p^0 = r^0 = b - Ax^0$. Then the iteration is

$$\begin{aligned} \alpha &= \frac{r^i \cdot r^i}{p^i \cdot Ap^i} \\ x^{i+1} &= x^i + \alpha p^i \\ r^{i+1} &= r^i - \alpha p^i \\ \beta &= \frac{r^{i+1} \cdot r^{i+1}}{r^i \cdot r^i} \\ p^{i+1} &= r^{i+1} + \beta p^i. \end{aligned}$$

There is only one MvM per step, and two dot products (the $r^i \cdot r^i$ can be saved from the previous iteration). MvM is easy to parallelize, but the dot products break parallel execution.

History of the conjugate gradient algorithm

"The method of conjugate gradients was developed independently by E. Stiefel of the Institute of Applied Mathematics at Zürich and by M. R. Hestenes with the cooperation of J. B. Rosser, G. Forsythe, and L. Paige of the Institute for Numerical Analysis, National Bureau of Standards. The present account was prepared jointly by M. R. Hestenes and E. Stiefel during the latter's stay at the National Bureau of Standards. The first papers on this method were given by E. Stiefel [1952] and by M. R. Hestenes [1951]. Reports on this method were given by E. Stiefel and J. B. Rosser at a Symposium on August 23-25, 1951. Recently, C. Lanczos [1952] developed a closely related routine based on his earlier paper on eigenvalue problem [1950]. Examples and numerical tests of the method have been by R. Hayes, U. Hoschstrasser, and M. Stein."

from Hestenes and Stiefel, 1952

The importance of technology

The CG was not devised earlier because

- CG does not work on slide rules
- CG has no advantage over Gauss elimination when using calculators
- CG has too much data exchange for a room full of human computers
- CG needs an appropriate computational engine

"The CG was discovered because Hestenes, Lanczos and Stiefel all had shiny, brand new toys (SWAC for Hestenes and Lanczos, Z4 for Stiefel)."

Dianne P. O'Leary, SIAM Linear Algebra Meeting, 2009.

References

- K. G. Wilson, *Confinement of Quarks*, *Phys. Rev.*, D10, 2445, 1974.
- M. Creutz, *Quarks, Gluons and Lattices*, Cambridge University Press.
- J. Smit, *Introduction to Quantum Fields on a Lattice*, Cambridge University Press.
- I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press.
- J. R. Shewchuk, *An Introduction to the Conjugate Gradient Method without the Agonizing Pain*, Penn State University Website (1994).