DESY 92 – 070 ISSN 0418 – 9833
May 1992

Effective Field Theories

G. Mack, T. Kalkreuter, G. Palma*, and M. Speh
II. Institut für Theoretische Physik, Universität Hamburg,
Luruper Chaussee 149, 2000 Hamburg 50

Abstract

Effective field theories encode the predictions of a quantum field theory at low energy. The effective theory has a fairly low ultraviolet cutoff. As a result, loop corrections are small, at least if the effective action contains a term which is quadratic in the fields, and physical predictions can be read straight from the effective Lagrangean.

Methods will be discussed how to compute an effective low energy action from a given fundamental action, either analytically or numerically, or by a combination of both methods. Basically, the idea is to integrate out the high frequency components of fields. This requires the choice of a "blockspin", i.e. the specification of a low frequency field as a function of the fundamental fields. These blockspins will be the fields of the effective field theory. The blockspin need not be a field of the same type as one of the fundamental fields, and it may be composite. Special features of blockspins in nonabelian gauge theories will be discussed in some detail.

In analytical work and in multigrid updating schemes one needs interpolation kernels $\mathcal{A}$ from coarse to fine grid in addition to the averaging kernels C which determines the blockspin. A neural net strategy for finding optimal kernels is presented.

Numerical methods are applicable to obtain actions of effective theories on lattices of finite volume. The special case of a "lattice" with a single site (the constraint effective potential) is of particular interest. In a Higgs model, the effective action reduces in this case to the free energy, considered as a function of a gauge covariant magnetization. Its shape determines the phase structure of the theory. Its loop expansion with and without gauge fields can be used to determine finite size corrections to numerical data.

*On leave of absence from Universidad de Santiago de Chile, Casilla 307, Correo 2, Santiago-Chile

Contents

1 What are Effective Field Theories?

We work in the context of Euclidean field theory.

- An effective field theory encodes the predictions of a quantum field theory at *low energies* .
- It is a *field theory* with a *low UV-cutoff* a_0^{-1} - of the order of the relevant particle masses. For instance it may live on a *lattice* of lattice spacing a_0.
- Ideally one should be able to read the physics straight from the *effective action* (= action of the effective theory) because loop corrections to tree amplitudes are small at sufficiently low UV-cutoff.

Problem to be discussed

Start from a given *fundamental action* (on the continuum or on a lattice of small lattice spacing $a \ll a_0$)

- How does one *define and compute the effective action*?

Sometimes effective actions are guessed on the basis of known properties.

Example: chiral theories (nonlinear σ-models) as effective theories of mesons. They embody the low energy theorems implied by PCAC.[1] These low energy theorems are sufficiently restrictive to fix the effective action at really low energy (i.e. low cutoff) up to a few coupling constants. By the definition of an effective theory, the question of the removal of the UV-cutoff does not pose itself. Nonrenormalizable σ-models are therefore acceptable.

But ideally, one would like not to guess the effective theory, but to justify it from the given fundamental theory.

Definition of the effective action in terms of the fundamental theory

The effective action will be a function of *low frequency fields* Φ - for instance, fields Φ on a lattice of lattice spacing a_0. They should be defined as *functions of the fundamental fields* φ,

$$\Phi = C(\varphi) \ ,$$

If C is linear, we write instead

$$\Phi = C\varphi \ .$$

C will be called the averaging operator and Φ is called *blockspin*. Later on we will consider fundamental theories which are gauge theories, and we will distinguish in notation between matter fields φ and gauge fields U. In general the low frequency matter fields will be functions not only of the matter fields φ but will also depend on the gauge fields,

$$\Phi = C(U, \varphi) \ \ , \text{ or } \quad \Phi = C(U)\varphi \ .$$

The definition of the effective action $\mathcal{H}_{eff}$ in terms of the fundamental Euclidean action $\mathcal{H}(\varphi)$ reads [2]

$$e^{-\mathcal{H}_{eff}(\Phi)} = \int \mathcal{D}\varphi \ \delta(\Phi - C(\varphi)) \ e^{-\mathcal{H}(\varphi)} \tag{1}$$

If the effective theory lives on a lattice Λ_0, then

$$\delta(\Phi - C(\varphi)) = \prod_\alpha \prod_{x \in \Lambda_0} \delta(\Phi_\alpha(x) - C_\alpha(\varphi)(x))$$

when Φ has components Φ_α. The 1-dimensional δ-function admits the integral representation

$$\delta(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \ e^{ik\xi} \ .$$

This formula, or (better) an approximation of the δ-function by a Gaussian

$$\delta_\beta(\xi) = \left(\frac{2\pi}{\beta}\right) e^{-\xi^2/2\beta} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \ e^{ik\xi} \ e^{-\beta k^2/2} \quad (\beta \ \text{ small})$$

can also be used when there are *Fermi fields.*

---

[1] PCAC = partially conserved axial vector current.

[2] Euclidean field theory may be regarded as a classical statistical mechanics. In this language, the action becomes the Hamiltonian. Therefore we use the letter $\mathcal{H}$.

Problems

- On what *number and kinds of fields* Φ_α should $\mathcal{H}_{eff}$ depend ?
- How to choose the *averaging operator* C.
  Averaging over space time regions will be involved in getting low frequency fields.
- How to *compute the functional integral* (1) by analytical or numerical means.
- How to *detect a bad choice of* C *or* Φ.

The answer to the last question is given by the requirement that *the effective action must have good locality properties.*

Let us emphasize that $\mathcal{H}_{eff}$ is *exact* in the sense that it yields exact expectation values for observables O which depend on the fundamental fields φ only through $\Phi_\alpha = C_\alpha(\varphi)$ so that $O = F(C(\varphi))$.

$$\begin{aligned} \langle O \rangle &= \frac{1}{Z} \int \mathcal{D}\varphi \ F(C(\varphi)) \ e^{-\mathcal{H}(\varphi)} \\ &= \frac{1}{Z} \int \mathcal{D}\Phi \ F(\Phi) \ e^{-\mathcal{H}_{eff}(\Phi)} \ . \end{aligned} \tag{2}$$

We illustrate the issues involved as well as the dangers of a bad choice of Φ or C at an example which is well understood.

Example: Discrete Gaussian model.

We consider first the case where the fundamental theory is a discrete Gaussian model in 3 dimensions. It lives on a lattice Λ of lattice spacing a. This model is the dual transform of a $U(1)$ lattice gauge theory in 3 dimensions. An exact computation of its effective action by convergent series expansions was used in [1] to give a rigorous proof that the 3-dimensional $U(1)$-lattice gauge theory shows linear confinement for all values of the gauge coupling constant g,with $\beta = 4\pi^2/g^2$. [3]

The points of the fundamental lattice will be denoted by letters z. The fundamental field is *integer valued* ,

$$n(z) \in \mathbb{Z} \ ,$$

and the fundamental action is [4]

$$\mathcal{H}(n) = \frac{1}{2\beta} \int_z [\nabla_\mu n(z)]^2 \ .$$

The difficult case is when β/a *is large* (i.e. weak gauge coupling g.) Note that β/a is dimensionless in 3 dimensions.

A choice was made to use a Pauli-Villars cutoff for the effective theory, so that the effective theory lives on the original lattice Λ, to begin with. The cutoff a_0^{-1} was chosen to be of the order of the ultimate physical mass m of the theory, in agreement with the philosophy of effective field theory. But it is essential to choose as the low frequency field a *real field*

$$\Phi(z) \ \text{real} \ .$$

[3] The convergent expansions are iterated Mayer expansions. They combine expansions at various length scales. Similar expansions can be written down which live on a multigrid [2] - cp. later.

[4] We use lattice notations as follows on a d-dimensional lattice Λ of lattice spacing a: $\int_z = a^d \sum_{z\in\Lambda}$ and $\nabla_\mu f(z) = a^{-1}[f(z+e_\mu) - f(z)]$, e_μ= lattice vector of length a in μ-direction.

This illustrates the point that *the fields Φ of the effective theory need not be fields of the same type as those in the fundamental theory* .

One finds that a local sine-Gordon action is an excellent approximation to the effective action when β is large

$$\mathcal{H}_{eff}(\Phi) = \int_z \left(\frac{1}{2\beta} \left[\nabla_\mu \Phi(z)\right]^2 + 2\mathsf{z} \left[1 - \cos\Phi(z)\right] \right) \tag{3}$$

with

$$\begin{aligned} \mathsf{z} &= a^{-3} \exp\left[-\beta v_{Cb}(0)\right] \quad \text{for large } \beta \ , \\ v_{Cb}(0) &= 0.2527 \ a^{-1} \ . \end{aligned} \tag{4}$$

$v_{Cb}(0)$ is the Coulomb potential at zero distance on the 3-dimensional lattice. (We chose not to write the terms in $\mathcal{H}_{eff}$ which implement the Pauli-Villars cutoff.)

We proceed to a discussion of the physics of the model on the basis of of the effective action.

1) **Mass:** The $[1 - \cos]$ potential has its minima at $\Phi = 2\pi\cdot$integer. We shall see in a moment that tunneling between different minima is very much suppressed, so that the global symmetry

$$\Phi(x) \mapsto \Phi(x) + 2\pi N \ , \quad N \text{ integer}$$

is spontaneously broken, and we may restrict attention to one of the minima, e.g. $\Phi = 0$, and expand around it. Set

$$\begin{aligned} \Phi &= \beta^{1/2}\, \Psi \quad , \\ m^2 &= \mathsf{z}\ 2\beta = 2\beta\ a^{-3} \exp\left[-\beta v_{Cb}(0)\right] \ . \end{aligned} \tag{5}$$

Then the effective action takes the form

$$\mathcal{H}_{eff} = \frac{1}{2} \int_z \left(\left[\nabla_\mu \Psi(z)\right]^2 + m^2\ \frac{1}{2}\ \Psi(z)^2 + [\cdots] \right) \ . \tag{6}$$

The nonlinear terms $[\cdots]$ are small for large β (provided Ψ is sufficiently near to the selected minimum so that the effective action is sufficiently small to give Ψ a nonnegligible probability). The effective theory is therefore very nearly a free field theory with a nonvanishing mass m which is given by (5). Numerical simulations confirm this [3].

The low lying excitations of the model are therefore spin waves with a finite mass which is exponentially small in units of the original inverse lattice spacing a^{-1}. This is easy to read off the effective action, but is not at all apparent from the fundamental action.

2) **Tunneling:** Because of the UV-cutoff we can qualitatively think of the theory as living on a lattice Λ_0 of lattice spacing $a_0 =$ const. $\cdot\, m^{-1}$. The effective action becomes

$$\mathcal{H}_{eff} = \sum_{x \in \Lambda_0} \frac{a_0^3}{2\beta} \left[\nabla_\mu \Phi(x)\right]^2 + 2\mathsf{z}_0 \left[1 - \cos\Phi(x)\right] \tag{7}$$

$$\mathsf{z}_0 = a_0^3 \mathsf{z} = \text{const.}'\ m^{-3} \mathsf{z} \tag{8}$$

$$= \text{const}''\beta^{-3/2} \exp\left[\beta v_{Cb}(0)/2\right] \tag{9}$$

We see that $\mathsf{z}_0 \mapsto \infty$ when $\beta \mapsto \infty$. Therefore the hills of the $2\mathsf{z}_0 [1 - \cos\Phi]$ - potential become very high when β becomes large. As a result, tunneling becomes very much suppressed.

The height and widths of the hills of the potential determine the surface tension of the discrete Gaussian model (the free energy per area of an interphase between domains where $\langle\Phi\rangle = 2\pi N$ with different N). This surface tension is equal to the string tension α of the dual $U(1)$-gauge theory. Treating the tunneling in a semiclassical approximation results in

$$\alpha = 8m\beta^{-1} \ . \tag{10}$$

Note that

$$\alpha/m^2 \mapsto \infty \quad \text{as } \beta \mapsto \infty \ .$$

Thus *there are two different physical mass scales in this theory.*

3) **2-dimensional discrete Gaussian model:**

Next we turn to the 2-*dimensional discrete Gaussian model.* This model is the dual transform of the XY-model (plane rotator) in 2 dimensions with Villain action.

A rigorous computation of the effective action of this model has not been reported in the literature, but the method used for the 3-dimensional model can be expected to carry over for large β. In 2 dimensions, β is dimensionless.

The effective action comes out the same as in 3 dimension, except for the appearance of different powers of a (for dimensional reasons) and the following difference in the expression for z. Expression (4) results from an approximation which is legitimate in 3 dimensions, but not in 2. Originally there stands the Yukawa potential $v_\mu(0)$ at zero distance in place of $v_{Cb}(0)$, with mass μ equal to the Pauli-Villars cutoff a_0^{-1}. We are interested in the limit where μa is very small. In 3 dimensions, the zero mass limit exists, and we may approximate $v_\mu(0)$ by $v_{Cb}(0)$. But in 2 dimensions, the zero mass limit does not exist. Instead, $v_\mu(0)$ increases logarithmically with $(\mu a)^{-1}$ as the cutoff $\mu = a_0^{-1}$ is lowered. As a result,

$$\mathsf{z}_0 = a_0^2\, \mathsf{z} = (a\mu)^{-2} \exp\left[-\beta v_\mu(0)/2\right] \tag{11}$$

decreases when the cutoff is lowered, if β is large enough, and the mass in tree approximation decreases faster than the cutoff μ. Therefore, the cutoff can be lowered indefinitely. Since z_0 decreases indefinitely, the effective action tends to a massless free field theory in the limit of very low cutoff, provided β is large enough. This is the Kosterlitz-Thouless phase of the model.

There is no particular reason to believe that the formula (7) for the effective action remains accurate for intermediate values of β. But the nature of the Kosterlitz-Thouless phase phase transition could be read off, if one had an accurate expression for the effective action at intermediate β.

Lessons from the model

- The physics can readily be read from the effective action, but is not visible in the fundamental action.
- A good choice of the blockspin Φ (real as opposed to integer) is crucial.

If we had chosen an integer blockspin, the spin waves of zero or small mass m as the dominant low energy excitations would not be represented by fields in the action. Moreover, one would meet a disaster while calculating the effective action. If one tried to compute it by lowering the cutoff step by step (which is what an iterated Mayer expansion does), the action would become nonlocal over distances m^{-1} while the UV-cutoff μ is still very much higher than m. A good choice of the blockspin is

$$\Phi(z) \approx \text{average of } n(w) \text{ over a cube of side } a_0$$

using a smoothened step function for the cube.

Stages of complexity of the effective action

To be specific, let us start from the fundamental action of a Higgs model with nonabelian gauge fields living on Λ. Λ may be either the continuum or a lattice of small lattice spacing. We seek an effective theory on a lattice Λ_0 of lattice spacing a_0.

We identify the points $x \in \Lambda_0$ with hypercubes in the fundamental lattice Λ.

These hypercubes are called blocks. Sometimes we will also regard Λ_0 as a sublattice of Λ, so that $x \in \Lambda_0$ fixes a site in Λ, e.g. the block center.

One may envisage different kinds of and approximations to the effective action which require different kinds of blockspins

1. Ignore both the gauge field U and the fermion fields ψ.
   Approximate the fundamental theory by a $\lambda\varphi^4$-theory. Choose the block average of the scalar field as blockspin [4]

   $$\Phi_\alpha(x) = \int_z C(x,z)\ \varphi_\alpha(z) = Z_0^{-1/2}(a_0)\ av_{z\in x}\varphi_\alpha(z)\ .$$

   with $C(x,z) \propto$ step function. A cutoff dependent factor $Z_0^{1/2}(a_0)$ (wave function renormalization) has been included with the aim to normalize the kinetic term in the effective action in the standard way.

2. Ignore the gauge field, but consider fermions.
   An effective action which has chiral symmetry requires $2^{d/2}$ flavours f in d dimensions (for even d). Otherwise the effective action will inevitably become nonlocal (Nielsen-Ninomiya theorem [5]). Given fundamental fermion fields with $2^{d/2}$ flavours, we may block to staggered fermions on Λ_0 of lattice spacing $\frac{1}{2}a_0$. Λ_0 decomposes into elementary cells of sidelength a_0 containing 2^d sites each. Only translations by integer multiples of a_0 are true translations. Sites within one elementary cell carry different pseudoflavour (a combination of spin and flavour). A suitable blockspin is of the form

   $$\Psi(x) = \int_z \sum_{f,a} C_a^f(x,z)\ \psi_a^f(z) \tag{12}$$

   where a, f are spinor and flavour indices. C_a^f averages over a hypercube of sidelength a_0 centered at x, and is only nonzero when (a,f) matches with the pseudoflavour of x. Details will be given later on.

3. Admit dynamical gauge fields.
   The blockspin for the Higgs field will take the form

   $$\Phi_\alpha(x) = \int_z C(U|x,z)\ \varphi_\alpha(z) \tag{13}$$

similarly as in 1., but with an averaging kernel C which depends on the gauge field U for reasons of covariance. A formula like (12) with U-dependent kernels C_a^f will define the blocked Fermi fields. In addition one will need to define blocked gauge fields. This will be discussed later on.

4. Composite Higgs.
There may be no fundamental Higgs field in the fundamental action. In this case one may try to define a composite Higgs field Φ in the effective action by a formula of the following kind
$$\Phi_\alpha(x) = \int_z \int_w \sum \overline{\psi}_a^f \ C_{ab}^{fg}(U|x, zw) \ \lambda_\alpha \psi_b^g(w) \ .$$
with matrices λ_α in colour space (colour indices of the Fermi fields ψ are omitted).

5. Fully effective theory.
In this case, the fields in the effective action are all interpolating fields for physical particles. They can be defined as composite fields similar to 4., but they will have to be gauge invariant. For instance, the physical electron field could take the form
$$\Psi(x) = \int_z \int_w \psi_a^f(z) \ C_a^f(U|x, zw) \ \varphi(w)$$
where ψ_a^f are the left and right handed fundamental electron field and φ is the fundamental Higgs field.

One may wish to compute the ultimate effective action in a sequence of steps, lowering the cutoff in stages. In this case, one may have to proceed to different kinds of definitions of blockspins at different values of the UV-cutoff corresponding to different length scales of compositeness.

2 The Constraint Effective Potential

Starting from a fundamental theory on a finite volume, the constraint effective potential is defined as the special case of the effective action in which the lattice Λ_0 consists of a single point. In this case, the block spins Φ may be vectors in colour or flavour space, but they have no space time argument x.

Consider for instance a fundamental theory which is a Higgs model with or without dynamical gauge fields. For simplicity, we do not include fermions. One introduces
$$\Phi = \text{(gauge covariant) magnetization}$$
by a formula similar as before. The constraint effective potential $V_{eff}(\Phi)$ was introduced in [6] and is defined by
$$e^{-V_{eff}(\Phi)} = \int \mathcal{D}\varphi \ \mathcal{D}U \ \delta(\Phi - C\varphi) \ e^{-\mathcal{H}(U,\varphi)} \tag{14}$$
and similarly in the absence of gauge fields. C depends on U in the presence of gauge fields.

$\exp\{-V_{eff}(\Phi)\}$ gives us the probability density of the magnetization Φ. But
$$V_{eff} \sim \ \text{volume} \quad .$$
Therefore the minima of V_{eff} become extremely sharp in the limit of large volume, and in the infinite volume limit, the only possible values which Φ may assume are absolute minima of

V_{eff}. Therefore, the shape of V_{eff} in the large volume limit gives information about the phase structure of the theory. One has to look for minima. If there are several absolute minima [5] which are transformed into each other by a symmetry, then that symmetry is spontaneously broken. For plots which show the crossover of the effective potential from a single shape (symmetric phase) to double well behaviour (broken phase) see Fig. 1.

If the height of different relative minima changes when one varies a parameter in the theory in such a way that the absolute minimum jumps, then the theory has a first order phase transition.

3 Renormalization Group Picture

In the renormalization group approach [7] one composes the blocking

$$\begin{aligned} \Lambda &\mapsto \Lambda_0 \\ a &\mapsto a_0 \\ \mathcal{H} &\mapsto \mathcal{H}_{eff} \end{aligned}$$

from a sequence of blockings, lowering the cutoff in steps

$$\begin{aligned} \Lambda &\equiv \Lambda_N \mapsto \Lambda_{N-1} \mapsto \ldots \mapsto \Lambda_1 \mapsto \Lambda_0 \\ a &\equiv a_N \mapsto a_{N-1} \mapsto \ldots \mapsto a_1 \mapsto a_0 \\ \mathcal{H} &\equiv \mathcal{H}^N \mapsto \mathcal{H}^{N-1} \mapsto \ldots \mapsto \mathcal{H}^1 \mapsto \mathcal{H}^0 \equiv \mathcal{H}_{eff} \end{aligned}$$

where $a_{j-1} = La_j$ with $L = 2, 3$ or so.

The theory with action $\mathcal{H}^j$ lives on a lattice of lattice spacing a_j. Considering one such theory at a time, one goes to the unit lattice by setting $a_j = 1$. In this way one obtains a sequence of actions $\widetilde{\mathcal{H}}^j$.

$$\widetilde{\mathcal{H}}^N \mapsto \widetilde{\mathcal{H}}^{N-1} \mapsto \ldots \mapsto \widetilde{\mathcal{H}}^1 \mapsto \widetilde{\mathcal{H}}^0 \equiv \widetilde{\mathcal{H}}_{eff} \tag{15}$$

These actions can be compared, since they all live on the same lattice with lattice spacing 1. This yields the

$$\text{renormalization group flow: } \widetilde{\mathcal{H}}^j \mapsto \widetilde{\mathcal{H}}^{j-1} \quad .$$

Imagine that the actions are somehow parametrized. One distinguishes between relevant, marginal, and irrelevant parameters. These parameters are also called "running coupling constants".

Relevant: A small change in a relevant coupling constant in $\mathcal{H}^j$ changes $\mathcal{H}^{j-n}$ very much for large n.

Irrelevant: A small change in an irrelevant coupling constant changes $\mathcal{H}^{j-n}$ arbitrarily little for large n.

Marginal: in between.

[5] $\mathcal{V}_{eff}$/volume becomes convex in the infinite volume limit because the height of maxima of V_{eff} between minima increases less fast than the volume.

Figure 1: Multigrid Monte Carlo data for $\lambda\varphi^4$-theory [15] with error bars as a function of the magnetization Φ for bare coupling $\lambda_u = 16.376$ and fit with the one loop formula for the constraint effective potential for three values of the bare mass a) $m_u^2 = -1.14$, b) $m_u^2 = -1.15$, c) $m_u^2 = -1.16$.

In the limit of small a, many renormalization group steps will be needed. As a result, knowledge of all relevant and marginal coupling constants in $\mathcal{H}$ suffices to determine $\mathcal{H}_{eff}$. Theories with actions $\mathcal{H}$ whose relevant and marginal parameters all agree are said to belong to the same universality class.

Let us remark that one can also lower an UV-cutoff (Pauli-Villars, for instance) infinitesimally in one step. In this case the renormalization group flow is determined by a functional differential equation known as the renormalization group differential equation [8].

In the conventional approach, all the actions $\mathcal{H}^j$ depend on the same type of fields. But we noted before that more flexibility should be left. It may be necessary to change the kind of fields at some length scales (of compositeness) a_j. In particular, $\mathcal{H}^j$ may depend on different kinds of fields than the fundamental action $\mathcal{H}$ when $j < N$.

Effective observables

In principle, expectation values $\langle O\rangle_{\mathcal{H}}$ of arbitrary observables in the fundamental theory may be translated into expectation values $\langle O_{eff}\rangle_{\mathcal{H}_{eff}}$ in the effective theory, such that

$$\langle O\rangle_{\mathcal{H}} = \langle O_{eff}\rangle_{\mathcal{H}_{eff}} \ .$$

But this requires that one computes effective observables in addition to the effective Hamiltonian [9],

$$O_{eff}(\Phi) = e^{\mathcal{H}_{eff}(\Phi)} \int \mathcal{D}\varphi \ O(\varphi) \ \delta(\Phi - C(\varphi)) \ e^{-\mathcal{H}(\varphi)} \quad . \tag{16}$$

4 Computation of Effective Actions by Perturbation Theory (Loop Expansion)

for linear averaging maps C in theories with or without gauge fields.

Scalar free field theory

$$\mathcal{H}_{\text{free}}(\varphi) = \frac{1}{2}\int_z \varphi(z)[-\Delta + m^2]\varphi(z) = \frac{1}{2}(\varphi, [-\Delta + m^2]\varphi) \ .$$

Herein, Δ may be either the ordinary Laplacian, or the covariant Laplacian in some given external gauge field, and the original theory may live either on the continuum or on a lattice Λ. Its sites are denoted by z, w. Our aim is to compute an effective action on a lattice Λ_0 of lattice spacing $a_0 \leq O(m^{-1})$. Sites in Λ_0 are denoted by x, y.

We assume a linear averaging map C,

$$\Phi(x) = \int_z C(x,z)\varphi(z) \equiv C\varphi(x) \ .$$

In the absence of a gauge field, we may choose

$$a_0^d \ C(z,x) = \begin{cases} 1 & \text{if } z \in x \\ 0 & \text{otherwise} \ . \end{cases} \tag{17}$$

The original lattice Λ of sites z is covered by a block lattice Λ_0, whose sites x are blocks with side length a_0, so we write $z \in x$ if the point z is in the block x. More generally, C may depend on an (external) gauge field.

Following Kupiainen and Gawedzki [4, 10], one splits the field φ into a low frequency part ψ, called the "background field", and a high frequency part ζ, called the "fluctuation field". The low frequency field is supposed to be obtainable from the block spin Φ on Λ_0 by interpolation with an interpolation kernel $\mathcal{A}$.

$$\begin{aligned} \varphi(z) &= \psi(z)+\zeta(z) \\ &= \mathcal{A}\Phi(z)+\zeta(z)\ , \quad \text{with} \quad \mathcal{A}\Phi(z)=\int_{x\in\Lambda_0}\mathcal{A}(z,x)\Phi(x)\ . \end{aligned} \tag{18}$$

If we demand that

$$C\mathcal{A}=\mathbb{1}\ , \tag{19}$$

then it follows that the fluctuation field has zero block average,

$$C\zeta=0\ . \tag{20}$$

Now we insert the field split into $\mathcal{H}_{\text{free}}$. Let us consider the case $m=0$. Generalization is obvious.

$$(\varphi,\Delta\varphi)=(\mathcal{A}\Phi,\Delta\mathcal{A}\Phi)+(\zeta,\Delta\zeta)+2(\zeta,\Delta\mathcal{A}\Phi)\ . \tag{21}$$

We demand that $\mathcal{A}$ is chosen in such a way that the mixed term vanishes,

$$(\zeta,\Delta\mathcal{A}\Phi)=0\ . \tag{22}$$

This gives

$$\mathcal{H}_{\text{free}}=\frac{1}{2}(\Phi,-\Delta_{eff}\Phi)+\frac{1}{2}(\zeta,\Delta\zeta)\ , \tag{23}$$

with

$$\Delta_{eff}=\mathcal{A}^*\Delta\mathcal{A}\ . \tag{24}$$

Since $C\zeta=0$ it follows that the mixed term (22) vanishes as desired, provided

$$-\Delta\mathcal{A}=C^*u^{-1} \tag{25}$$

for some u^{-1}.

Applying $\mathcal{A}^*$ to both sides of (25) it follows that

$$u^{-1}=-\Delta_{eff}\ . \tag{26}$$

The operator Δ_{eff} acts according to

$$\Delta_{eff}f(x)=\int_{y\in\Lambda_0}\Delta_{eff}(x,y)f(y) \tag{27}$$

with kernel

$$\Delta_{eff}(y,x)=\int_z\mathcal{A}(z,y)^\dagger\Delta\mathcal{A}(z,x). \tag{28}$$

Here, Δ acts on the argument z, and † denotes the hermitean conjugate of a matrix, or the transpose, if the matrix is real.

(25) becomes

$$\Delta\mathcal{A}(z,x)=\int_y C(y,z)^\dagger\Delta_{eff}(y,x)\ . \tag{29}$$

If $C(x,z)$ is the step function (17) then (29) says that $\Delta\mathcal{A}(z,x)$ should be constant on blocks y as a function of z. The kernels $\mathcal{A}$ and Δ_{eff} are determined as solutions of (29) together with the subsidiary condition (19).

Remark: The field $\psi(z) = \mathcal{A}\Phi(z)$ minimizes $\mathcal{H}_{\text{free}}(\psi)$ subject to the subsidiary condition that ψ has the prescribed block average Φ, viz.

$$C\psi(x) = \Phi(x) \ .$$

This remark can be used to compute the kernel $\mathcal{A}$ by standard optimization algorithms.

Numerical results confirm that the kernel $\mathcal{A}(z,x)$ decays exponentially with the distance of z from the center of the block x, with decay length a_0. This is true with and without an external gauge field in Δ, see for instance Fig. 5 of Sect. 8. As a result, $\Delta_{eff}(x,y)$ decays exponentially as a function of $|x-y|$ with decay length one lattice spacing a_0.

Now we are ready to compute the effective action. By definition

$$e^{-\mathcal{H}_{eff}(\Phi)} = \int \mathcal{D}\varphi\delta(\Phi - C\varphi)e^{-\mathcal{H}_{\text{free}}(\varphi)} \ . \tag{30}$$

We shift the field φ by an amount depending on Φ, using $\zeta = \varphi - \mathcal{A}\Phi$ as a new variable of integration. Since $C\mathcal{A} = \mathbb{1}$ by assumption, it follows that

$$e^{-\mathcal{H}_{eff}(\Phi)} = e^{\frac{1}{2}(\Phi,\Delta_{eff}\Phi)} \int \mathcal{D}\zeta \ \delta(C\zeta) \ e^{-\mathcal{H}_{\text{free}}(\zeta)} \ . \tag{31}$$

The ζ-integral merely produces a constant independent of Φ. Therefore

$$\mathcal{H}_{eff}(\Phi) = \frac{1}{2}(\Phi, -\Delta_{eff}\Phi) + \text{ const.} \quad . \tag{32}$$

$\mathcal{H}_{eff}$ has good locality properties because of the aforementioned decay properties of $\Delta_{eff}(x,y)$. There are exponential tails, but this is tolerable and in general inevitable.

Split of the propagator

The split of the field φ induces a split of the free propagator $v = (-\Delta)^{-1}$ into a low frequency propagator v_{low} and a high frequency propagator (fluctuation field propagator) Γ. For proper choice of the normalization factor $\mathcal{N}$,

$$d\mu_\Gamma(\zeta) = \mathcal{N}^{-1}\mathcal{D}\zeta \ e^{-(\zeta,-\Delta\zeta)/2} \ \delta(C\zeta) \tag{33}$$

is a normalized Gaussian measure with covariance

$$\Gamma(z,w) = \mathcal{N}^{-1} \int \mathcal{D}\zeta \ \delta(C\zeta) \ e^{-(\zeta,-\Delta\zeta)/2} \ \zeta(z) \ \zeta(w) \ .$$

Thinking of the δ-function as a limit of a Gaussian, one sees that

$$\Gamma = \lim_{\kappa\mapsto\infty} (-\Delta + \kappa C^* C)^{-1} \tag{34}$$

Since ζ is supposed to be the high frequency part of the field φ, it is a field with an infrared cutoff (equal to the UV-cutoff of the effective theory). Therefore one expects that $\Gamma(z,w)$ decays exponentially with the distance of z from w, with decay length a_0. Analytical and numerical results confirm that this is indeed the case. This remains true when there is a a gauge field in Δ, provided the kernel C is chosen in the right way. A proper choice will be introduced later.

Let us now compute the 2-point function.

$$
\begin{aligned}
\langle\varphi(z)\varphi(w)\rangle &\propto \int \mathcal{D}\Phi\ \mathcal{D}\varphi\ \delta(\Phi - C\varphi)\ e^{-\mathcal{H}_{\text{free}}(\varphi)}\ \varphi(z)\ \varphi(w) && (35)\\
&= \int \mathcal{D}\Phi\ \mathcal{D}\zeta\ e^{(\Phi,-\Delta_{eff}\Phi)+(\zeta,-\Delta\zeta)}\ \delta(C\zeta)(\mathcal{A}\Phi(z)+\zeta(z))(\mathcal{A}\Phi(w)+\zeta(w)) && (36)
\end{aligned}
$$

Since $u = (-\Delta_{eff})^{-1}$ it follows that

$$
\mathcal{N}_\Phi \cdot \int \mathcal{D}\Phi\ e^{(\Phi,-\Delta_{eff}\Phi)/2}\ \Phi(x)\ \Phi(y) = u(x,y)
$$

for a proper choice of the normalization factor $\mathcal{N}_\Phi$.
As a result we get the following final result for $\langle\varphi(z)\varphi(w)\rangle \equiv v(z,w)$

$$
v(z,w) = v_{low}(z,w) + \Gamma(z,w)\ , \tag{37}
$$

with

$$
v_{low} = \mathcal{A}\ u\ \mathcal{A}^*\quad ,\quad u = -\Delta_{eff}^{-1}\ . \tag{38}
$$

More explicitly, (37) reads

$$
\begin{aligned}
\Gamma(z,w) &= v(z,w) - \int\int_{x,y\in\Lambda^1} \mathcal{A}(z,x)u(x,y)\mathcal{A}^\dagger(w,y) && (39)\\
v &= -\Delta^{-1}\quad \text{and}\quad u = -\Delta_{eff}^{-1} && (40)
\end{aligned}
$$

Perturbation theory

We will now show how to compute effective actions by perturbation theory. To be specific, let us consider an action for a scalar field φ of the form

$$
\mathcal{H}(\varphi) = \int_z \{-\frac{1}{2}\varphi(z)\Delta\varphi(z) + \mathcal{V}(\varphi(z))\} = (\mathcal{H}_{\text{free}} + V)(\varphi)\ . \tag{41}
$$

with

$$
\mathcal{V}(\xi) = \frac{1}{2}m_u\ \xi^2 + \frac{1}{4!}\ \lambda_u\ \xi^4\ , \tag{42}
$$

for instance. Introduce

$$
V_\Phi(\zeta) = \int_z \mathcal{V}(\mathcal{A}\Phi(z) + \zeta(z))\ . \tag{43}
$$

This determines a self interaction of the high frequency field ζ which depends parametrically on the block spin Φ.

Proceeding as before one finds the following expression for the effective action

$$
\mathcal{H}_{eff}(\Phi) = -\frac{1}{2}(\Phi, \Delta_{eff}\Phi) + \widehat{V}_{eff}(\Phi) \tag{44}
$$

with

$$
e^{-\widehat{V}_{eff}(\Phi)} = \int d\mu_\Gamma\ e^{-V_\Phi(\zeta)}\ . \tag{45}
$$

We remember that $d\mu_\Gamma$ is the normalized free field measure with propagator Γ. The logarithm of the ζ-integral can be computed by standard perturbation theory. It is given by a sum of Feynman diagrams, whose lines represent propagators Γ, and whose vertices depend parametrically on Φ.

The resulting effective action $\mathcal{H}_{eff}$ will have good locality properties, provided the fluctuation field propagator $\Gamma(z,w)$ and the interpolation kernel $\mathcal{A}(x,z)$ decay exponentially with decay length a_0.

Perturbation expansions of the kind described here were first done by Kupiainen and Gawedzki [4, 10], as part of a rigorous renormalization group analysis. Constraint effective potentials were also computed by another method in [11].

Constraint effective potential

We specialize the results of the previous subsection to the case of the constraint effective potential. In this case, Λ_0 has only one point x. Therefore the kernel of Δ_{eff} can only have (x,x) as its only pair of variables. We admit the possibility that Δ is the covariant Laplacian in an external gauge field. Of course, Δ_{eff} will also depend on U in this case. We abbreviate

$$\Delta_{eff}(x,x) = \epsilon_0 \ .$$

Eq. (44) simplifies to

$$\mathcal{H}_{eff}(\Phi) \equiv V_{eff}(\Phi) = \Phi\ \epsilon_0\ \Phi + \widehat{V}_{eff}(\Phi) \ . \tag{46}$$

with $\widehat{V}_{eff}$ from (45).

In the presence of an external gauge field U, ϵ_0 depends on U. $\widehat{V}_{eff}$ will also depend on U, because the high frequency propagator Γ does, and the kernel $\mathcal{A}$ as well. In addition there is a gauge field dependence in the normalization factors of the Gaussian measures. This leads to a Φ-independent, but U-dependent term in $\widehat{V}_{eff}$ which will be ignored for now, but which has to be taken into account later on when we consider dynamical gauge fields.

Without gauge field, $\epsilon_0 = 0$. But in the presence of a gauge field which is not a pure gauge, $-\Delta$ is strictly positive, and therefore ϵ_0 is also strictly positive.

Let us emphasize that $\mathcal{H}_{eff}$ has no infrared problems because Γ has a built-in infrared cutoff.

1-loop approximation

In 1-loop approximation, the contribution $\widehat{V}_{eff}$ to the constraint effective potential (46) becomes [12, 13, 14]

$$\widehat{V}_{eff}(\Phi) = \int\int_{z,w} \mathcal{V}'(\mathcal{A}\Phi(z))\Gamma(z,w)\mathcal{V}'(\mathcal{A}\Phi(w)) + \frac{1}{2}\mathrm{Tr}\ \ln(-\Delta + \mathcal{V}''(\mathcal{A}\Phi)) \tag{47}$$

where a prime ($'$) indicates a derivative with respect to the argument, and $\mathcal{V}''$ is the multiplication operator defined by

$$(\mathcal{V}''(\mathcal{A}\Phi)f)(z) = \mathcal{V}''(\mathcal{A}\Phi(z))f(z) \ . \tag{48}$$

In graphical notation

$$\mathrm{Tr}\ \ln(-\Delta + \mathcal{V}''(\mathcal{A}\Phi)) \quad = \quad \sum_{n\geq 1} \frac{(-)^n}{n} \int_{z_1} \cdots \int_{z_n}$$

with Feynman rules

$$z \;\text{- - - - - - - -}\; w = \Gamma(z,w) \quad , \qquad\qquad \otimes_z = \mathcal{V}''(\mathcal{A}\Phi(z)) \ .$$

The first term in expression (47) is not 1-particle irreducible. This term vanishes in the absence of gauge fields, and more generally it vanishes up to second order in the gauge coupling constant.

The contribution $\widehat{V}_{eff}$ to the full effective action (44) on a lattice Λ_0 of more than one point has the same form. The 1-particle reducible contribution is important in this case [10].

Use of the 1-loop formula to compute finite size corrections to the constraint effective potential

For definiteness, let us consider $\lambda\varphi^4$-theory.

From the constraint effective potential/volume in the infinite volume limit one can extract renormalized coupling constants, provided one knows the wave function renormalization constant from an independent calculation.

In fact, according to the standard renormalization of the $\lambda\varphi^4$-theory, one needs to define the three quantities m_r, λ_r and Z_φ (renormalized mass, self-coupling constant and wave function renormalization constant respectively). This can be achieved in the symmetric phase through the second and the fourth derivatives of the constraint effective potential at its minimum, viz.

$$m_r^2 = \frac{U''_{eff}(\Phi_{\min})}{Z_\varphi^{-1}} \tag{49}$$

and for the self-coupling constant

$$\lambda_r = \frac{U''''_{eff}(\Phi_{\min})}{Z_\varphi^{-2}}. \tag{50}$$

In the broken phase it is useful to define the renormalized square mass also through (49) but the self-coupling constant λ_r is now defined through the tree level relation:

$$\lambda_r = 3\, m_r^2 \,/ v_r^2 \tag{51}$$

where v_r is the vacuum expectation value of the field $\varphi_r = Z_\varphi^{-1/2}\varphi$ and $U_{eff} = V_{eff}/\text{volume}$.

One can obtain a renormalized loop expansion for the constraint effective potential by manipulating the bare expansion. Eqs. (49) - (51) determine the renormalized coupling constants as functions of the bare parameters. Inverting this relations, one obtains the bare quantities as a function of the renormalized ones. One may use them to express $\mathcal{V}$ in expression (47) (or its corresponding n-loop generalization) for the constraint effective potential as a function of the renormalized coupling constants. Reordering the resulting expansion in number of loops gives the desired expansion in terms of λ_r, m_r.

In practice, numerical computation of constraint effective potentials is only feasible on finite and not too large lattices.

The following strategy, which was first proposed by one of us (G.P.) in [12], eliminates the need to use very large lattices. It makes it possible to extract renormalized coupling constants from numerical data for values of the bare coupling constants which are very close to the critical line, and for lattice sizes which need not be large compared to the correlation length.

Triviality of $\lambda\varphi^4$-theory implies that the renormalized coupling constant λ_r is necessarily small when one is close to a critical point, provided the UV-cutoff is high enough, i.e. provided $a \ll$ correlation length.

Table 1: Renormalized parameters from the fits of Fig. 1 and the corresponding values obtained from Lüscher and Weisz (^{LW}).

	$m_r(\infty)$	$\lambda_r(\infty)$	m_r^{LW}	λ_r^{LW}
$m_0^2 = -1.14$	0.079 ± 0.003	7.5 ± 0.6	0.08	7.4 ± 0.2
$m_0^2 = -1.15$	0.06 ± 0.005	7.1 ± 0.1	0.06	7.2 ± 0.2
$m_0^2 = -1.16$	0.13 ± 0.008	7.8 ± 0.1	0.13	7.6 ± 0.2

Therefore, a perturbative calculation of the constraint effective potential on a *finite lattice* as a function of the renormalized coupling constants will yield accurate results. In the practical application [12, 13] it was verified that the two loop corrections are indeed very small.

The result of this computation is fitted to numerical data taken at various values of the bare coupling constants to extract the renormalized coupling constants as functions of the bare ones.

Some results of this kind are shown below. Numerical results of Meyer and Mack [15] are shown in Fig. 1 together with a fit by a renormalized 1-loop formula for the constraint effective potential. Numerical results for the constraint effective potential for 4-component φ^4-theory at stronger coupling but further away from the critical line were presented in [16].

The results are compared with analytical results of Lüscher and Weisz [17] in Table 1 .

Perfect agreement is found. In this way, the nonperturbative part of the analytical computation of Lüscher and Weisz is sucessfully tested.

5 Averaging and Interpolation Kernels C and $\mathcal{A}$ in the Presence of Gauge Fields, for Bosons

In the presence of a gauge field U, the definition of a blockspin Φ for a (n-component) scalar field φ and the split of this field into low frequency and high frequency part read

$$\begin{aligned} \Phi &= C(U)\varphi \ , \\ \varphi &= \mathcal{A}(U)\Phi + \zeta \ , \\ C(U)\mathcal{A}(U) &= \mathbb{1} \ . \end{aligned} \tag{52}$$

The operators C and $\mathcal{A}$ have kernels

$$\mathcal{A}(U)\Phi(z) = \int_{x\in\Lambda_0} \mathcal{A}(U|z,x)\ \Phi(x)\ , \tag{53}$$

$$C(U)\varphi(z) = \int_{z\in\Lambda} C(U|x,z)\ \varphi(z)\ . \tag{54}$$

As usual we denote points of the block lattice Λ^0 by x, y and points of the fundamental lattice Λ by z, w.

In the absence of gauge fields, a good choice of C is given by

$$C^*(z,x) \equiv C(x,z)^\dagger = \begin{cases} 1 a_0^{-d} & \text{if } z \in x \\ 0 & \text{otherwise} \end{cases} ,$$

(† denotes the hermitean conjugate of a matrix.) This satisfies

$$- \Delta_{N,x} C^*(z,x) = \begin{cases} \epsilon_0(x)\ C^*(z,x) & \text{if } z \in x \\ 0 & \text{otherwise} \end{cases} . \tag{55}$$

$\Lambda_{N,x}$ is the lattice version of the Laplacian with Neumann boundary conditions on the boundary of the block x, and $\epsilon_0(x) = 0$ is its lowest eigenvalue. In the presence of a gauge field, the Laplacian with Neumann boundary conditions takes the form

$$\Delta_{N,x} f(z) = a^{-2} \sum_{w\ nn.z, w \in x} [U(z,w) f(w) - f(z)] \ . \tag{56}$$

$U(z,w)$ is the lattice gauge field attached to the link from w to z, the summation runs over those nearest neighbours w of z which are inside the block x.

Proposal: [18] Retain the eigenvalue equation (55) for C in the presence of gauge fields, with covariant Laplacian $\Delta_{N,x}$ as defined in (56), and $\epsilon_0(x)$ equal to its lowest eigenvalue. Of course, both C and $\epsilon_0(x)$ will depend on the gauge field.

The interpolation kernel $\mathcal{A}$ may also be defined by the same equation as before. There should be $u(y,x)$ such that

$$- \Delta \mathcal{A}(z,x) = \int C^*(z,y)\ u(y,x) \ . \tag{57}$$

Here Δ is the gauge covariant Laplacian on the whole lattice. u is a Lagrange multiplier which is determined by the subsidiary condition (52).

In the special case when the block lattice Λ_0 consists of a single point x, the Neumann boundary conditions are inoperative, because the block x has no boundary when periodic boundary conditions for Λ are in force. Therefore $\Delta_{N,x} = \Delta$. The equation for $\mathcal{A}$ is therefore satisfied by

$$\mathcal{A} = C^* \quad \text{if } \Lambda_0 = \text{ single site} \quad . \tag{58}$$

Let us next discuss the arbitrariness in the solution of (55). Consider for example the gauge group $G = SU(2)$. For fixed x, the 2×2 matrix $C(z,x)$ may be thought to be composed of two 2-vectors, each of which is a solution of the eigenvalue equation (55). (It can be proven that there exist two linearly independent solutions with the same eigenvalue.) Arbitrary linear combinations of these two solutions will also be solutions. This reflects itself in the arbitrariness

$$C^*(z,x) \mapsto C^*(z,x)\ \Omega(x) \tag{59}$$

of the 2×2-matrix solution C of (55), where $\Omega(x)$ are arbitrary 2×2-matrices. To eliminate this arbitrariness, we may impose subsidiary conditions

1. Normalization: $CC^* = \mathbb{1}$.
   This restricts the arbitrariness to unitary matrices $\Omega(x)$.

2. Gauge transformation property:
   Select a distinguished site $\hat{x}$ in the block x – for instance the center of the block – and demand that $C^*(\hat{x}, x) > 0$ as a matrix. This makes the kernel C unique. It will transform under a gauge transformation [6] according to

$$C^*(z,x) \mapsto g(z) C^*(z,x) g(\hat{x})^{-1} \ . \tag{60}$$

[6] Remember that under a gauge transformation $\varphi(z) \mapsto g(z)\varphi(z)$ and $U(z,w) \mapsto g(z)U(z,w)g(w)^{-1}$. The covariant Laplacian $\Delta_{N,x}$ is gauge covariant. Therefore the eigenvalue equation (55) is also gauge covariant.

It follows that the blockspin $\Phi(x)$ transforms like a field sitting at the block center $\widehat{x}$,

$$\Phi(x) \mapsto g(\widehat{x})\Phi(x) \ . \tag{61}$$

Given C, the kernel $\mathcal{A}$ is also unique and it transforms under gauge transformations exactly like C^*.

There exists an efficient algorithm for computing the gauge field dependent kernels C and $\mathcal{A}$ [18].

Let us consider the implications for the effective potential in the presence of dynamical gauge fields, (14). Since the U-dependent kernel C will depend on the choice of $\widehat{x}$, the gauge covariant magnetization will also depend on $\widehat{x}$. But it is readily verified that the effective potential (14) is independent of the choice of $\widehat{x}$ [14].

6 Effective Potential for a Higgs Model with Dynamical Gauge Fields

Let us consider a Higgs model with dynamical gauge fields but without fermions, with a gauge group $SU(2)$ and a 2-component scalar field φ. Let

$$\Phi = \text{gauge covariant magnetization} = \int_z C(z)\varphi(z) \ , \tag{62}$$

defined with the help of the gauge covariant kernel $C(x,z)$ which was described in the last section. The argument x is redundant because there will be only one point x in Λ_0. Of course, C and Φ will depend on the gauge field U, and so does $\mathcal{A} = C^*$. By definition,

$$\begin{aligned} e^{-V_{eff}(\Phi)} &= \int \mathcal{D}U \ e^{-\widetilde{V}(U,\Phi)} \\ e^{-\widetilde{V}(U,\Phi)} &= \int \mathcal{D}\varphi \ \delta(\Phi - C\varphi) \ e^{-\mathcal{H}(U,\varphi)} \ . \end{aligned} \tag{63}$$

$\mathcal{H}$ is the standard action for a Higgs model. Using an appropriate lattice definition of the field strength tensor $F_{\mu\nu}$ [19], the Wilson action for the lattice gauge field can be written in standard $F_{\mu\nu}F_{\mu\nu}$-form, and

$$\begin{aligned} \mathcal{H} &= \int_z \left[\frac{1}{4}F_{\mu\nu}F^*_{\mu\nu} + \frac{1}{2}\langle\varphi, -\Delta\varphi\rangle + \mathcal{V}(\varphi(z))\right] \\ \mathcal{V}(\varphi(z)) &= \frac{1}{2}m_u^2\varphi^2(z) + \frac{1}{4!}\lambda_u\varphi^4(z) \ . \end{aligned} \tag{64}$$

Δ is the gauge covariant Laplacian. It depends on U.

The functional integral will be performed in two steps. (For details of the loop expansion, see [14].) Calculations by a different method were reported in [20] for a $U(1)$-gauge field.)

φ integration

$\mathcal{D}\varphi \ \delta(\Phi - \int_z C(z)\varphi(z))$ is an unnormalized Gaussian measure with nonzero mean. First we shift the field to bring the mean to zero. We introduce a new integration variable ζ in place of φ,

$$\varphi(z) = \eta(z) + \zeta(z) \quad \text{with} \quad \eta(z) = \mathcal{A}(z) \ \Phi \tag{65}$$

Since Λ_0 is a single point, the choice $\mathcal{A} = C^*$ is in agreement with previous definitions. The interpolation kernel satisfies therefore the eigenvalue equation

$$- \Delta \mathcal{A}(z) = \epsilon_0 \, \mathcal{A}(z) \ , \tag{66}$$

ϵ_0 = lowest eigenvalue, Δ = covariant Laplacian on the lattice with periodic boundary conditions.

Using $\mathcal{A}^* \mathcal{A} = C \mathcal{A} = \mathbb{1}$, the kinetic term becomes

$$\langle \varphi, \ -\Delta \varphi \rangle = \Phi \epsilon_0 \Phi + \langle \zeta, -\Delta \zeta \rangle \tag{67}$$

so that

$$e^{-\widetilde{V}(U,\Phi)} = e^{-\frac{1}{2}\Phi \epsilon_0(U) \Phi} \ \int \mathcal{D}\zeta \ \delta(\int_z C(z)\zeta(z)) \quad \exp \left(-\frac{1}{2} \langle \zeta, -\Delta \zeta \rangle + \int_z \mathcal{V}(\eta(z) + \zeta(z) \right) \ . \tag{68}$$

The eigenvalue problem (66) can be solved by time independent perturbation theory which is familiar from elementary quantum mechanics. Set

$$U(z + e_\mu, z) = 1 - g A_\mu(z) + \ldots \ . \tag{69}$$

To order g^2 one finds

$$\epsilon_0 = \frac{g^2}{8} \int_z |A_\mu(z)|^2 + \frac{g^2}{4} \sum_{\mu,\nu} \int_z \int_w A_\mu(z) \Gamma(z, w) A_\nu(w) \ . \tag{70}$$

The high frequency propagator Γ is only needed to zeroth order in g and the expression valid in a scalar theory without gauge fields can therefore be used for it at this point.

In conclusion

$$\int \left\{ \frac{1}{4} F_{\mu\nu} F^*_{\mu\nu} + \frac{1}{2} \Phi \epsilon_0 \Phi \right\} = \int \frac{1}{4} F_{\mu\nu} F^*_{\mu\nu} + \frac{1}{2} m^2(\Phi) \int |A_\mu|^2 + \ldots . \tag{71}$$

with

$$m^2(\Phi) = \frac{1}{4} g^2 |\Phi|^2 \ . \tag{72}$$

We see that a Φ-dependent mass term for the gauge field has appeared.

In the full theory, the fluctuations of the random variable Φ tend to zero in the infinite volume limit (cp. J. Klauders lectures at this school [21]), and its value becomes fixed to the position $\widehat{\Phi}$ of an absolute minimum of $V_{eff}(\Phi)$. As a result

$$m^2(\widehat{\Phi}) = \text{W-meson mass} \ . \tag{73}$$

The integral

$$\int \mathcal{D}\zeta \exp \left(-\frac{1}{2} \langle \zeta, -\Delta \zeta \rangle + \int \mathcal{V}(\eta(z) + \zeta(z)) \right) \tag{74}$$

is computed by expanding the exponent in the gauge coupling constant g and using the standard formula for a Gaussian integral

$$\int \mathcal{D}\zeta \exp \left\{ -\sum_z \left\{ \frac{1}{2} \langle \zeta | M_z \zeta \rangle + \langle \zeta | J_z \rangle \right\} \right\} = (\det M)^{1/2} \exp \left(\frac{1}{2} \sum_z \langle J_z | M_z^{-1} J_z \rangle \right) \ . \tag{75}$$

Expressing it in terms of the *normalized* Gaussian measure $d\mu_\Gamma$ with covariance (= free propagator) Γ, the integral (74) equals

$$e^{-\frac{1}{2}\mathrm{Tr}_C \ \ln\Gamma} \int d\mu_\Gamma(\zeta)\ e^{-\int_z \mathcal{V}(\eta(z)+\zeta(z))}\ . \tag{76}$$

The high frequency propagator Γ depends on the gauge field, and so does $\mathcal{A}=C^*$ and therefore η, defined in (65). Γ was defined in (34) and

$$\mathrm{Tr}_C = \text{trace on the space of functions } f \text{ which obey } Cf=0\ .$$

With some effort, $\mathrm{Tr}_C\ \ln\Gamma$ can be evaluated to second order in g. It is a quadratic expression in the gauge fields which will add to the kinetic term of the gauge field.

Gauge field integration

Expanding everything to second order in the gauge coupling constant g, the integration over the gauge fields becomes a Gaussian integration which may be performed. Thanks to term $\Phi\epsilon_0\Phi$ in in (67) and also additional quadratic terms in the vector potential A_μ which resulted from the φ integration, the kinetic term for the gauge field is nondegenerate.

In fact, it turns out that the whole quadratic contribution in the gauge field A_μ in the large volume limit is equal to

$$S_g^q[A]=\frac{1}{2v}\sum_k\sum_{\mu\nu}\delta_{rs}\left\{\delta_{\mu\nu}[\hat{k}^2+m^2(\Phi)]-\left(1+\frac{m^2(\Phi)}{\hat{k}^2}(\Phi)\Theta\left(\hat{k}^2\right)\right)\hat{k}_\mu\,\hat{k}_\nu\right\}|\tilde{A}_\mu(k)|^2 \tag{77}$$

where $\tilde{A}_\mu(k)$ is the Fourier transform of the gauge field $A_\mu(z)$ and $\Theta(\hat{k}^2)$ is equal to one for $\hat{k}^2\neq 0$ and the interpolation to continuous $\hat{k}^2$ vanishes faster than $\hat{k}^2$ for $\hat{k}^2\to 0$.

The quadratic operator of (77) *does* have an inverse (the propagator) and therefore no gauge fixing is necessary!

The final result in finite volume is given in (120) in the Appendix.

Comparison with known results [22] for the conventional effective potential reveals that (120) simplifies in the infinite volume limit to

$$V_{eff}(\Phi)/\text{ volume } = \text{conventional effective potential in the Landau gauge.} \tag{78}$$

The conventional effective potential is the Legendre transform of the generating function for Greens functions, with a constant field as its argument. It has been computed before [22] on a finite lattice L^4 in the Landau gauge.

If one performs a formal continuum limit in addition to the infinite volume limit in order to replace sums over k by integrals, the expression (120) simplifies (after inserting counterterms into the action) to the following expression for the one-loop constraint effective potential $U_{eff}=V_{eff}/$volume [14].

$$\begin{aligned} U_{eff}(\Phi) = \ & U_{cl}(\Phi)+\tfrac{1}{64\pi^2}\left\{(U''_{cl}(\Phi))^2\left[\ln(U''_{cl}(\Phi))-\tfrac{1}{2}\right]\right.\\ & \left.+3\left(\tfrac{U'_{cl}}{\Phi}(\Phi)\right)^2\left[\ln\left(\tfrac{U'_{cl}(\Phi)}{\Phi}\right)-\tfrac{1}{2}\right]+9\mu_W^4(\Phi)\left[\ln(\mu_W^2(\Phi))-\tfrac{1}{2}\right]\right\}\end{aligned} \tag{79}$$

with

Figure 2: The Constraint Effective Potential in the presence of gauge fields in the infinite and large cutoff limit. The gauge coupling was fixed to be $g_r^2 = 0.4$, the renormalized mass m_r^2 is given by $m_r^2 = -\frac{2}{3}\frac{\lambda_r}{g_r^2}$ and the scalar self coupling constant assumes values on both sides of the critical line (solid line). We observe that the minimum jumps from zero to $\Phi_{\min}$ when the parameters are varied from the symmetric to the Higgs phase. This means that the system has a weak first order transition.

$$U_{cl}(\Phi) = \frac{1}{2}\, m_r^2\, \Phi^2 + \frac{1}{4!}\, \lambda_r\, \Phi^4 \quad \text{and} \quad m^2(\Phi) = \frac{1}{4}\, g_r^2\, \Phi^2. \tag{80}$$

The expression (79) is manifest gauge invariant and it agrees with the known result of Anna Hasenfratz and Peter Hasenfratz [23] for the conventional effective potential computed in the Landau gauge. Φ is the unrenormalized covariant magnetization, i.e. no wave function renormalization factor is included in its definition.

Small renormalized couplings $\lambda_r = O(g_r^4)$ are within the domain of validity of the perturbative result of (79). Its shape is shown in Fig. 2. Inspection of Fig. 2 shows that the absolute minimum jumps when the parameters are varied. Therefore there is a first order phase transition at weak couplings. This was noted before by Hasenfratz and Hasenfratz [23], following the pioneering work of Coleman and Weinberg [24] .

7 Blockspin for Fermions

As already discussed in T. DeGrand's lecture [25], the naive discretization of the continuum fermion action results in a lattice model with unphysical low energy modes, called "doublers". To reduce the number of flavours, i.e. the degeneracy of the lattice Dirac operator, one can use one-component "staggered" fields which live on different lattice sites. In d dimensions, a staggered fermion field describes physical fermions with $2^{[d/2]}$ flavours.

We shall first be concerned with blockspins of the form (12), i.e. we ignore gauge fields for now and consider only fermions.

Figure 3: Part of a pseudoflavour lattice Λ^H for $d = 2$, $H \in \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$, embedded into $\Lambda = \mathbb{Z} \times \mathbb{Z}$.

Free Staggered Fermions on the Lattice

For simplicity let us first look at free staggered fermions on a d-dimensional lattice $\Lambda_{a/2} = (\frac{1}{2}a\mathbb{Z})^d$ of lattice spacing $a/2$. Only translations by integer multiples $\vec{n}a$ of a are regarded as true lattice translations, i.e. survivors of translations in the continuum. These true translations map the sublattice $\Lambda_a = (a\mathbb{Z})^d$ of lattice spacing a into itself.

We imagine that the lattice $\Lambda_{a/2}$ is divided into hypercubes of sidelength a containing 2^d sites each. Sites within one hypercube are distinguished by their "pseudoflavour" H. H is specified by a set of distinct indices μ_i

$$H = \{\mu_1\mu_2\ldots\mu_h, 0 \leq h < d, \mu_1 < \mu_2 < \ldots < \mu_h\} \quad . \tag{81}$$

Sites x_H in $\Lambda_{a/2}$ with the same pseudoflavour H are of the form

$$x_H \equiv \hat{x} + \frac{1}{2}e_H, \quad e_H = \sum_{\mu\in H} e_\mu, \quad \hat{x} \in \Lambda_a$$

where e_μ is the lattice vector of length a in μ-direction in Λ_a. The sites x_H form a sublattice Λ_a^H of $\Lambda_{a/2}$ (see Fig. 3), viz. $\Lambda_a^H = \Lambda_a + \frac{1}{2}e_H$.

Pseudoflavour is a combination of spin and flavour in the sense that there is an expansion

$$\sum_H \chi(x_H)\; dx^H = \sum_{\alpha,f} \Phi_\alpha{}^f(x)\; Z_\alpha{}^f \quad . \tag{82}$$

which exhibits the field $\chi(x_H)$ at site x_H in a hypercube as a linear combination of field components $\Phi_\alpha{}^f$ with spinor index α and flavour index f ($Z = (Z_\alpha{}^f)$ is a matrix of differential forms). Each index takes values $1, \ldots, 2^{[d/2]}$. For details, see [26].

Blocking Consistent with the Symmetry

We distinguish between two kinds of blockings: from the continuum to some lattice, and from one lattice to a coarser lattice. We wish to block in such a way that as much is preserved of internal and space time symmetries as is possible.

Blocking from the Continuum to the Lattice.

Given Fermi fields $\Psi^f_\alpha(z)$ on the continuum with $2^{[d/2]}$ flavours f, let $\chi^H(z)$ be the linear combination of field components Ψ^f_α of given pseudoflavour. The staggered blockspin on the lattice $\Lambda_{a/2}$ has the form

$$\Xi(x) = \int_z C^H(x,z)\ \chi^H(z) \quad \text{if } x \in \Lambda^H_a \subset \Lambda_{a/2} \quad . \tag{83}$$

Here $\int_z$ is an integral over continuous space.

In the absence of gauge fields the kernel of the averaging operator is given by

$$(\frac{a}{2})^d\ C^H(x,z) = \begin{cases} 1 & \text{if } x \in \Lambda^H_a, \max\ |x^\mu - z^\mu| \le \frac{a}{2} \\ 0 & \text{otherwise} \ . \end{cases} \tag{84}$$

i.e. one averages over a cube of sidelength a with center x. Note that the cubes around block lattice sites x, y may overlap if x, y have different pseudoflavour (see below). It can be verified that this choice of blockspin leads to good locality properties of the effective free action on the lattice [27].

We proceed with considerations on the staggered fermion symmetry group.

Residual Symmetry on the Lattice.

For bosons, any flavour symmetry G in the continuum is preserved on the lattice; continuous space time translations get broken to lattice translations and similarly for Euclidean rotations.

For staggered fermions, the situation is different [28]: blocking as described above from the continuum, the total symmetry $\mathcal{G} = G \times \mathcal{T}$ (e.g. flavor group $G = SU(2)$ for $d = 4$, translation group $\mathcal{T}$) gets broken to a subgroup $\mathcal{G}_a$ on a lattice of lattice spacing $a/2$. $\mathcal{G}_a$ is generated by translations e_μ by integer multiples of a and by certain combinations εd^H of flavour transformations ($\varepsilon = \pm 1$), and translations by $e_H/2$. The action T_a of the lattice symmetry group on a staggered field on the lattice $\Lambda_{a/2}$ reads as follows

$$(T_a(\varepsilon d^K)\chi)(y) = \varepsilon\ \check{\rho}_{H,K}\ \chi(y + \frac{1}{2}e_K) \quad \text{if } y \in \Lambda^H_a \ . \tag{85}$$

The symbols $\check{\rho}_{H,K}$ denote sign factors as in the Clifford product of forms $\vee$, defined by

$$dx^H \vee dx^K = dx^\mu \wedge dx^\nu + \delta^{\mu\nu} \qquad \text{and} \qquad dx^H \vee dx^K = \check{\rho}_{H,K} dx^{H\triangle K} \ ,$$

with the symmetric difference $H \triangle K = (H \cup K)\backslash(H \cap K)$ of sets H, K of pseudoflavour indices.

This agrees with the known symmetry group of standard staggered fermion actions [29]. Note that translations by e_H are true translations.

Fine-to-Coarse Blocking.

The fine-to-coarse blocking step maps fields which live on the fine lattice $\Lambda_{a/2}$ onto fields living on the coarser lattice $\Lambda_{L_b a/2}$ of lattice spacing $L_b a/2$, with an integer scale factor $L_b > 1$. It is natural to demand that the symmetry group $\mathcal{G}_{L_b a}$ on the coarse lattice should be a subgroup of the symmetry group $\mathcal{G}_a$ on the finer lattice

$$\mathcal{G}_{L_b a} \subset \mathcal{G}_a \quad . \tag{86}$$

In addition to the true translations by $e'_H = L_b e_H$ there will be combinations [7]

$$\varepsilon d'_H \in \mathcal{G}_{L_b a}$$

of flavor transformations and translations by $\frac{1}{2}e'_H$.

The symmetry operations d'_H must be a combination of some d_K and a true translation e. From (85), it can be seen that d'_K contains an admixture of translation by $-e'_K/2$. Since d_K contains an admixture of a translation by $\frac{1}{2}e_K$, we must have

$$-\frac{1}{2}e'_H \quad = \quad -\frac{1}{2}e_K + e \quad \text{with} \quad e = \sum_\mu n_\mu e_\mu \quad . \tag{87}$$

Since $e'_H = L_b e_H$, the only way how this can happen is that $H = K$ and $e = \frac{1}{2}(L_b - 1)e_H =$ *true translation*. This requires that L_b is odd, hence $L_b = 3, 5, \ldots$ but *not* $L_b = 2$!

This result might have been expected, since only for odd L_b the pseudoflavour (and parity) on the fine lattice $\Lambda_{a/2}$ and on the block lattice $\Lambda_{L_b a/2}$ match nicely when we regard $\Lambda_{L_b a/2}$ as sublattice of $\Lambda_{a/2}$ (see Fig. 4 below).

For a true symmetry it should not matter in which order we apply the blocking step and the symmetry operations: the blocking map C should commute with the action T of the symmetry under $\mathcal{G}_{L_b a} \subset \mathcal{G}_a$, viz.

$$T_{La}(d'_H)\; C(x,z) = C(x,z)\; T_a(d_H)\; T_a(\frac{1}{2}\,(L_b - 1)\, e_H) \quad . \tag{88}$$

This is fulfilled if C is translation invariant so that

$$C(x + e'_H, z + L_b e_H) = C(x,z) \tag{89}$$

and if in addition C has the support property

$$C(x,z) = 0 \quad \text{unless } x, z \text{ carry the same pseudoflavour.} \tag{90}$$

In words, the support property states the following. If $y \in \Lambda^H_{L_b a/2}$ then the (staggered) blockspin $\Xi(y)$ is a weighted sum of fields $\chi(z)$ at sites with the same pseudoflavour H. Blockings with such a property were considered for 2-dimensional $U(1)$ gauge fields by Ben Av et al. [30]. However, they chose a scale factor $L_b = 2$.

[7] A prime ($'$) refers to the coarser lattice.

Generalization to Nontrivial Gauge Fields

We now admit dynamical gauge fields in our staggered fermion theory; the averaging kernel C will then have to depend on the gauge field for reasons of covariance as in (13). Our interest will gradually shift towards a practical implementation of the renormalization group picture outlined in Sect. 3, known as the *multigrid approach* to numerical computations in lattice field theories [8].

To specify a blocking prescription, we retain blocks as introduced in Sect. 5: the solution of the gauge covariant eigenvalue equation

$$-\Delta_{b.c.} C^*(z,x) = \lambda_0(x) C^*(z,x) \tag{91}$$

yields a gauge covariant kernel C. Results reported below show that (91) produces a good blockspin when $\Delta_{b.c.}$ is the Laplace operator with appropriate boundary conditions (basically Neumann b.c.). The covariant Laplace operator connects only lattice sites of the same pseudoflavour; therefore (91) is consistent with the condition (90) above.

In a multigrid approach, one needs interpolating kernels $\mathcal{A}$ from coarse to fine lattices in addition to the averaging kernels C. The kernels $\mathcal{A}$ should be smooth. In the present context, there are a priori two possible definitions of smoothness.

1. "Laplacian" smoothness.

χ possesses Laplacian smoothness if

$$(\nabla_\mu \chi, \nabla_\mu \chi) \equiv (\chi, -\Delta\chi)$$

is small, with the *covariant Laplacian* Δ: for our model of staggered fermions this is expected to work if the field strength tensor $F_{\mu\nu}$, conveniently defined on the lattice [19] by

$$F_{\mu\nu}(z) = U(z + \frac{1}{2}e_\mu, \nu)\ U(z,\mu) - U(z + \frac{1}{2}e_\nu, \mu)\ U(z,\nu)\ . \tag{92}$$

is small. This is the standard definition of smoothness.

The problem is that there are no smooth functions in this sense if the gauge fields are disordered, because Δ is strictly positive in this case.

2. "Diracian" smoothness.

χ possesses Diracian smoothness if

$$(\not{\nabla}\chi, \not{\nabla}\chi) \equiv (\chi, -\not{\nabla}^2\chi)$$

is small. Since

$$-\not{\nabla}^2 = -\Delta + \sigma_{\mu\nu}F_{\mu\nu}\ , \tag{93}$$

sites of different pseudoflavour are connected.

When one wants to compute propagators of fermions by use of the multigrid approach, then one should in principle use interpolating kernels $\mathcal{A}$ which possess Diracian smoothness. Unfortunately this is impractical for reasons of storage space. Laplacian smooth kernels $\mathcal{A}$ are much more practical. Possible compromises are discussed in [26].

In Sect. 9 the problem of smoothness of the interpolation kernel $\mathcal{A}$ (see Sect. 5) will be readdressed using optimization wisdom borrowed from neural computation.

---

[8] Those readers who are not familiar with multigrid methods may first want to read Sect. 8; a more detailed mathematical account can be found in [31].

Figure 4: The good choice of blocks to average over. $C(z,x)$ is only nonvanishing if site z has the same pseudoflavour (symbol) as the block-center $\hat{x}$ (encircled symbol). Therefore the seemingly overlapping blocks (square) have actually no sites in common.

Special Choice which yields a Good Blockspin

In the language of Sect. 5, we know that we have made a good choice of blockspin if the effective action remains *local* after arbitrarily many or arbitrarily big blocking steps. The averaging kernel $C(x,z)$ which is nonvanishing only if z and x have the same pseudoflavour, and which satisfies the Laplace equation (90) with Neumann boundary conditions on block boundaries, has proven to be good in that sense. This result extends to blocking from continuum action with $2^{[d/2]}$ flavours to staggered fermion field on the lattice. Details will be given in the next section. Fig. 4 visualizes the overlapping of blocks which do not share sites of the same pseudoflavour as the block centers. In numerical investigations [32] non-overlapping blocks were used as well. But then one does not retain the successful pure gauge limit.

A successful block spin requires $2^{[d/2]}$ flavors. There is no consistent way of putting a single chiral particle on the lattice (this is the statement of the theorem by Nielsen and Ninomiya [5]): Therefore there is no good blockspin either for a single chiral field, as we had already remarked in Sect. 1.

8 Efficient Computation of Gauge Covariant Propagators

A quenched simulation of lattice QCD requires the evaluation of products of Dirac propagators $S(U)$. S depends on the actual gauge field configuration U. In a Monte Carlo simulation of full QCD, one needs also the fermion determinant

$$\det{}^{-1/2} S(U) = \mathcal{N}^{-1} \int \mathcal{D}\phi \ \exp\left[-(\phi, S(U)\phi)\right] \ .$$

Since for every new gauge field configuration U, $S(U)$ needs to be recomputed or updated, it is important to have an efficient method to do so. The conjugate gradient (CG) or the minimal residual algorithm is state of the art. Great hopes to do better are attached to multigrid (MG) methods [30, 32, 33, 34, 35, 36, 37, 38].

Deterministic Multigrid Methods

For convenience, we will use the opposite convention in enumerating the lattices of a multigrid in this section. The fundamental (i.e. finest) lattice is denoted by $\Lambda = \Lambda^0$. There will be a sequence of lattices Λ^0, Λ^1, $\Lambda^2, \ldots$ of increasing lattice spacing a_i, viz. $a_{i+1} = L_b a_i$ with $a_0 = a$. We wish to solve an inhomogeneous linear equation

$$D_0 \chi_0 = f_0 \tag{94}$$

on the fundamental lattice $\Lambda = \Lambda^0$, for given f_0. In our case, $D_0 = -\nabla^2 + m^2$. After some relaxation sweeps on Λ^0 one gets an approximation $\tilde{\chi}_0$ to χ_0 which differs from the exact solution by an error $e_0 = \chi_0 - \tilde{\chi}_0$. The fundamental idea of the MG to the solution of partial differential (or difference) equations [31, 39] is that the error e_0 should become smooth very fast, although it may become small very slowly because of critical slowing down (CSD). The error satisfies the equation

$$D_0 e_0 = r_0 \tag{95}$$

which involves the residual $r_0 = f_0 - D_0 \tilde{\chi}_0$. If e_0 is smooth, it is determined to a very good accuracy by a function e_1 on the next coarser lattice Λ^1, and can be represented in the form

$$e_0 = \mathcal{A} e_1 \tag{96}$$

with an interpolation map $\mathcal{A}$ which should be so chosen that it maps functions on Λ^1 into smooth functions on Λ^0. Conversely, e_1 can be obtained from e_0 with the help of an averaging map C which satisfies

$$C\mathcal{A} = \mathbb{1} \ . \tag{97}$$

It follows that $e_1 = Ce_0$. Inserting (96) into (95) and acting on the result with C, we see that e_1 will satisfy the equation

$$D_1 e_1 = r_1 \tag{98}$$

with

$$D_1 = C D_0 \mathcal{A} \quad , \quad r_1 = C r_0 \ . \tag{99}$$

The problem has been reduced to an equation on the coarser lattice. If there is still too much CSD at this level, one may repeat the procedure, going to coarser and coarser lattices.

Given $\mathcal{A}$, a possible choice of C which satisfies (97) is $C = (\mathcal{A}^* \mathcal{A})^{-1} \mathcal{A}^*$. The kernel $\mathcal{A}(z, x)$ should be a smooth function of z. The appropriate notion of smoothness depends on the dynamics, i. e. on D_0, in general. *Smooth means little contributions from eigenfunctions to high eigenvalues of* D_0*.* This insight is confirmed by results reported below. The point is important in systems in gauge fields and for other disordered systems.

Gauge Covariance

In a gauge theory, kernels $\mathcal{A}$ and C should be chosen in a gauge covariant fashion; see the discussion in Sect. 5. Under gauge transformations on Λ^0 they transform according to

$$\begin{aligned} \mathcal{A}(z,x) &\mapsto g(z)\mathcal{A}(z,x)g(\widehat{x})^{-1} \ , \\ C(x,z) &\mapsto g(\widehat{x})C(x,z)g(z)^{-1} \ . \end{aligned} \tag{100}$$

This is consistent with $C\mathcal{A} = \mathbb{1}$.

The most general expression of a kernel $\mathcal{A}$ with this covariance property is a weighted sum of parallel transporters $U(\mathcal{C})$ along paths $\mathcal{C}$ from $\widehat{x}$ to z, e. g.

$$\mathcal{A}(z,x) = \sum_{\mathcal{C}:\widehat{x}\mapsto z} \varrho(\mathcal{C})\ U(\mathcal{C}) \ , \tag{101}$$

where $\varrho(\mathcal{C})$ are numbers. And analogously for C. We will prefer not to specify the weights ϱ explicitly, but to determine $\mathcal{A}$, C as solutions of covariant equations.

Ground-State Projection Multigrid

In the ground-state projection MG method the averaging operator C from a grid to the next coarser grid is a projector on the ground-state of a local Hamiltonian. The adjoint of C satisfies a gauge covariant eigenvalue equation (cf. (55), (91)),

$$(-\Delta_{\mathrm{b.c.}}C^*)(z,x) = \epsilon_0(x)C^*(z,x) \ . \tag{102}$$

Remember that the kernel $C^*(z,x)$ which solves (102) is an $N_c \times N_c$ matrix which is in general not an element of the gauge group. (N_c is the number of colours.)

For bosons one may choose Neumann boundary conditions on block boundaries, $-\Delta_{\mathrm{b.c.}} = -\Delta_{N,x}$, as discussed in Sect. 5, (55). The lowest eigenvalue $\epsilon_0(x)$ of $-\Delta_{\mathrm{b.c.}}$ is a gauge invariant quantity which is a measure of disorder. A normalization condition $CC^* = \mathbb{1}$ may be imposed, and the covariance condition $C^*(\widehat{x},x) > 0$ (as a matrix) makes the definition of C unique. Kernels defined by (102) and $C^*(\widehat{x},x) > 0$ enjoy gauge covariance (100).

For staggered fermions, the Laplacian should be regarded as living on the sublattices which consist of sites with the same pseudoflavor. It connects only sites with the same pseudoflavor. Neumann boundary conditions on the Laplacian on the sublattice of sites z with the same pseudoflavor as x must be supplemented by the requirement that $C(x,z) = 0$ unless z and x have the same pseudoflavor. This extra requirement could be regarded as a consequence of supplementary boundary conditions.

There exists an efficient algorithm for computing C [18]. The programs are fully vectorized. In 4-dimensional $SU(2)$ gauge fields, computation of C on the whole lattice costs CPU time of the order of one heatbath sweep for the gauge field. Thus, it is not too expensive to solve (102), and it is not necessary to sacrifice gauge covariance by working with gauge field independent kernels in gauge-fixed U-configurations. (Gauge fixing costs also CPU time.)

Interpolation Kernel $\mathcal{A}$

Given the averaging kernel C, there exists an ideal choice of the interpolation kernel $\mathcal{A}$. It is determined as follows. For every function ("blockspin") Φ on Λ^1, $\phi = \mathcal{A}\Phi$ minimizes the action $\mathcal{H} = (\phi, D_0\phi)$ subject to the constraint $C\phi = \Phi$. Recalling the definition of smoothness, this

Figure 5: Optimal interpolation kernel $\mathcal{A}(z,y)$ for bosons on an 18^4 lattice in a quenched $SU(2)$ gauge field equilibrated at $\beta = 2.7$. ($L_b = 3$, periodic b. c.) Shown is a two-dimensional cut through the block center $y = 0$, z_3 and z_4 are fixed. The vertical axis gives the trace norm of $\mathcal{A}(z,y)$. (Remember that $\mathcal{A}(z,y) \in \mathbb{R} \cdot SU(2)$.)

can be rephrased: $\phi = \mathcal{A}\Phi$ is the smoothest possible function with prescribed block average $\Phi = C\phi$. It follows that $\mathcal{A}$ satisfies (106) below. With this choice of $\mathcal{A}$, the action $\mathcal{H}$ completely decouples into a sum of actions for the different MG layers. Moreover, $D_1 = (CC^*)\mathcal{A}^* D_0 \mathcal{A}$; this is selfadjoint for our C-kernels which satisfy $CC^* = 1\!\!1$.

A good "choice of blockspin", i.e. of C, is characterized by the fact that the ideal kernel $\mathcal{A}(z,x)$ ($z \in \Lambda^0$, $x \in \Lambda^1$) associated with it has good locality properties. This means that $\mathcal{A}(z,x)$ is big for $z \in x$, and decays exponentially in $|z - \widehat{x}|$ with decay length 1 block lattice spacing a_1 (cf. Sect. 4). Numerical computations of the ideal kernel $\mathcal{A}$ were done in quenched 4-dimensional $SU(2)$ gauge fields at various values of $\beta = 4/g^2$, including the case of a completely random gauge field ($\beta = 0$). (For bosons $D_0 = -\Delta + m^2$, while for fermions $D_0 = -\nabla\!\!\!/^2 + m^2$ with $m^2 = 0$ or small). It was verified that for *any* β-value the definition of the averaging kernel C as a solution of a gauge covariant Laplace eigenvalue equation (55), (91) (with Neumann boundary conditions on block boundaries) yields a good choice of blockspin in the sense described above. This is true both for bosons and for staggered fermions; see Figs. 5, 8. $\mathcal{A}$ decays exponentially over distance a_1 (nearly) as fast in the presence of gauge fields as in their absence.

Because of the exponential decay of the optimal kernel $\mathcal{A}$, one may try to approximate it by a simpler choice which fulfills $\mathcal{A}(z,x) = 0$ for $z \notin$ a certain neighborhood of block x. This neighborhood might be taken to be only the block x itself. The coarse grid operator would then be defined through $D_1 = \mathcal{A}^* D_0 \mathcal{A}$.

It follows from the optimization criteria discussed below that $\mathcal{A}$ should have the property that the lowest eigenmode ϕ_0 of D_0 should admit a representation of the form $\phi_0 = \mathcal{A}\Phi$ with a suitable Φ to a very good accuracy. The proposal of [18] to define $\mathcal{A}$ through the eigenvalue equation (102) with $-\Delta_{\text{b.c.}}$ taken to be a local Hamiltonian with Dirichlet instead of Neumann boundary conditions does not meet this requirement.

For periodic boundary conditions the Galerkin choice (also called variational coarsening) $\mathcal{A} = C^*$ is better. In this case D_1 can be defined as $D_1 = \delta^{-1} C D_0 C^*$ [38]. The fluctuating length of its matrix elements can be adjusted to fluctuate around the "right" value (of the exact $D_1(x,y)$) by tuning the real parameter δ.

Numerical Results for Propagators Computed by MG

Ground-state projection MG computations for propagators were done in 2-d $U(1)$ [34, 35] and (approximate ground-state projections) in 2-d and 4-d $U(1)$ and 2-d $SU(2)$ gauge fields [34, 36, 32]. The first ground-state projection MG computation of gauge covariant propagators in 4-d non-Abelian gauge fields (for gauge group $SU(2)$) was presented in Ref. [38]. Besides ground-state projection there is another approach for choosing the averaging kernel C. This approach is called "parallel transported MG". It is based on the representation (101) and is studied in Refs. [30].

Bosonic Propagators in 4-d $SU(2)$ Gauge Fields.

For bosons, $D_0 = -\Delta + m^2$. The MG method is of interest near criticality, i. e. for slowly decaying propagators. For nontrivial gauge fields we enforce slow decay by choosing m^2 negative and very close to the negative of the lowest eigenvalue $-m^2_{\text{cr}} > 0$ of $-\Delta$.[9] The relaxation time τ – defined by the *asymptotic* exponential decay of the norm of the residual or error – behaves like $\tau \propto (\triangle m^2)^{-z/2}$ for small $\triangle m^2 = m^2 - m^2_{\text{cr}}$.

Results in 4-dimensional $SU(2)$ gauge fields equilibrated with Wilson's action at various values of β [38]:

- The MG algorithm with the ideal interpolation kernel $\mathcal{A}$ eliminates CSD for *any* value of the gauge coupling; see Table 2. These computations were the first without CSD in non-trivial gauge fields. They prove that *ground-state projection is a good choice of* C, and that *the MG method can cope with the "frustration" (disorder) which is inherent in non-Abelian gauge fields.* However, the optimal $\mathcal{A}(z,y)$ is not translational invariant (except for $U \equiv \mathbb{1}$) and has support on all sites z of the fine lattice, for all sites y of the block lattice. Therefore the use of this kernel for production runs is impractical, but it was important to answer questions of principle. In particular this result confirms that there exists an appropriate notion of smoothness in disordered cases, as described above. The ideal $\mathcal{A}$ is the smoothest kernel which obeys $C\mathcal{A} = \mathbb{1}$.

- The most practical MG algorithm with $\mathcal{A} = C^*$ does not eliminate CSD in non-trivial gauge fields, $z = 2$ remains (as in 1-grid relaxation), although $z = 0$ in pure gauges; see Fig. 6. However, MG is able to outperform the conjugate gradient algorithm; see Fig. 7.

[9] Relaxation times of conventional iterative algorithms depend only on m^2 and *not* on the lattice size $|\Lambda|$; this is contrarily to the Dirichlet problem, for instance. In case of periodic b. c. there is only an implicit (slight) dependence on $|\Lambda|$ through the value of m^2_{cr}.

Table 2: Results of the idealized MG algorithm with lexicographic SOR (relaxation parameter ω) on 9^4 lattices Λ^0. Kernel $\mathcal{A}$ is the solution of the equation $([-\Delta + m_{\rm cr}^2 + \kappa\, C^* C]\mathcal{A})(z,y) = \kappa\, C^*(z,y)$ for large κ. This expression tends to the ideal $\mathcal{A}$-kernel described in the text for large κ.

β	$m_{\rm cr}^2$	optimal ω	τ for $\triangle m^2 \leq 10^{-3}$
∞	0	1.27	1.6
2.7	-0.8210607	1.38	1.9
2.5	-0.9477085	1.40	1.9
2.2	-1.2218471	1.45	1.9
1.8	-1.7567164	1.57	2.5
0	-2.7480401	1.69	5.2

Figure 6: Computation of bosonic propagators $(-\Delta + m_{\rm cr}^2 + \triangle m^2)^{-1}$. Relaxation times τ (in comparable work units) of iterative algorithms on an 18^4 lattice in a quenched $SU(2)$ gauge field at $\beta = 2.7$ with $m_{\rm cr}^2 = -0.7554339$. (MG with $\mathcal{A} = C^*$; Jac means Jacobi relaxation, while SOR stands for successive over-relaxation.)

- After the publication of Ref. [38], a practical nonlinear MG algorithm has been developed which eliminates the appearance of CSD when the mass is lowered, some volume dependence of the correlation times still remains, though. Details will be reported elsewhere [40].

Propagators for Staggered Fermions in 4-d $SU(2)$ Gauge Fields.

Of the two different proposals for the averaging kernel C made in [26], only the Laplacian choice has been implemented numerically yet. There, $-\Delta_{\rm b.c.} = -\Delta_{N,x}$ in (102), where Δ is the fermionic 2-link Laplacian defined in (93). The optimal interpolation kernel $\mathcal{A}$ associated with this C has the desired falloff properties in non-trivial gauge fields; see Fig. 8. Thus, the Laplacian choice for C defines a good blockspin for staggered fermions, and therefore it seems

Figure 7: Computation of bosonic propagators $(-\Delta + m_{\rm cr}^2 + \triangle m^2)^{-1}$. Convergence on an 18^4 lattice in quenched $SU(2)$ gauge fields at (a) $\beta = 10.0$, and (b) $\beta = 2.7$, with $m_{\rm cr}^2 = -0.1533739$, resp. -0.7554339. (MG with $\mathcal{A} = C^*$.)

not to be necessary to implement the more intricate proposals mentioned in [26]. However, numerical simulations with the optimal $\mathcal{A}$ have not been performed yet.

MG with variational coarsening $\mathcal{A} = C^*$ works, but up to now we had to restrict ourselves to relatively small lattices (up to 18^4) where this MG algorithm is not competitive with conjugate gradient. The authors of Refs. [30] report that their algorithm performs considerably better when they re-unitarize the blocked gauge field (matrix elements of the effective Dirac operator) such that it is again an element of the gauge group. However, our results indicate that the difference between using a unitary or a dielectric gauge field [19] on coarser layers can be compensated by adjusting the real parameter δ in D_1. By tuning δ, one can re-unitarize the blocked gauge field "on the average".

We are optimistic that the implementation of the nonlinear MG algorithm found for bosons (see above) also gives a practical and competitive MG method for fermions. This implementation is currently investigated. For details refer to [40].

9 Neural optimization of deterministic multigrids

The results of Sect. 8 confirm that the multigrid approach is able to handle inhomogeneous wave equations in disordered media, e.g. in the presence of gauge fields, but the appropriate notion of smoothness depends on the dynamics, i.e. on the operator D_0 in the equation which one wishes to solve. The interpolation kernels $\mathcal{A}$ should be smooth. Smooth means little contributions from eigenfunctions to high eigenvalues of D_0.

This means that the appropriate kernels $\mathcal{A}$ are dynamically determined and need to be

Figure 8: Optimal interpolation kernel $\mathcal{A}(z,y)$ – associated with the "Laplacian" choice for C – for staggered fermions on an 18^4 lattice in a quenched $SU(2)$ gauge field equilibrated at $\beta = 2.7$. ($L_b = 3$, periodic b. c.) $\mathcal{A}$ is the solution of the equation $\left([-\nabla\!\!\!\!/^2 + \kappa\, C^*C]\mathcal{A}\right)(z,y) = \kappa\, C^*(z,y)$ for large κ. Shown is a two-dimensional cut through the block center $y = 0$, z_3 and z_4 are fixed. The vertical axis gives the trace norm of $\mathcal{A}(z,y)$. (Remember that $\mathcal{A}(z,y) \in \mathbb{R} \cdot SU(2)$.) Note the support properties of $\mathcal{A}$, it has support on all even lattice sites and is vanishing on all odd sites. This is due to the non-vanishing field strength at finite β. In the limiting case of pure gauges the field strength term $\sigma_{\mu\nu}F_{\mu\nu}$ vanishes and the support of $\mathcal{A}$ is reduced to 1/16-th of $|\Lambda|$. Then we have 16 bosonic $\mathcal{A}$-kernels on 9^4 sublattices.

found before or together with the solution of the equation. (The averaging kernel C is often related to $\mathcal{A}$.) The question is how. There are interesting potential applications other than gauge theories [10]. Therefore a general strategy is desired.

Criteria for optimality

One may write down a cost functional whose minimum yields the optimal kernels, i.e. the fastest possible convergence of the multigrid iteration.

For definitenes and simplicity, think of a twogrid. Assume that updating consists of standard damped relaxation sweeps on the fundamental lattice to smoothen the errors, followed by an updating step $\chi_0 \mapsto \chi_0 - \mathcal{A}e_1$ which involves the exact solution e_1 of (98) on the coarse grid Λ^1, with some D_1 which is to be determined together with $\mathcal{A}$ and C.

Convergence speed is estimated by the norm of the iteration matrix ρ. The iteration matrix is the product $\rho = \rho_0\rho_1$ of the iteration matrix ρ_0 for the smoothing, and the iteration matrix ρ_1 for the coarse grid updating. The norm

$$\|\rho\| \leq \|D_0\rho_0\| \cdot \|D_0^{-1}\rho_1\| \ .$$

Only the second factor depends on $\mathcal{A}$, C and D_1. Inclusion of the operators $D_0^{\pm 1}$ in the factors reflects the philosophy that smoothing suppresses the high frequency modes of the error [31]. Choosing the trace norm, one is lead to determine $\mathcal{A}$, C and D_1 from the minimality condition for $E_1 = \|D_0^{-1}\rho_1\|^2$,

$$\begin{aligned} \min \quad &= \quad E_1 \ = \int\!\!\int_{z,w\in\Lambda^0} |\Gamma(z,w)|^2 && (103)\\ \Gamma(z,w) \quad &= \quad v(z,w) - \int\!\!\int_{x,y\in\Lambda^1} \mathcal{A}(z,x)\ \tilde{u}(x,y)\ C(y,w) && (104)\\ v \quad &= \quad D_0^{-1} \quad \text{and} \quad \tilde{u} = D_1^{-1} && (105)\end{aligned}$$

The minimality condition leads to conditions on kernels $\mathcal{A}$, C and on D_1. Suppose first that no constraints of practicality are imposed on the kernels. We assume that D_0 is selfadjoint. Optimizing $\mathcal{A}$, for given C and $\tilde{u} = D_1^{-1}$, one finds the condition $\Gamma C^* = 0$. That is,

$$D_0\mathcal{A} = C^* u^{-1} \tag{106}$$

with $u = \tilde{u}CC^*$. This reproduces the familiar condition for an ideal $\mathcal{A}$-kernel. Variation of the cost functional with respect to C yields $\mathcal{A}^*\Gamma = 0$. This is equivalent to a second condition which is obtained from (106) by interchange of the role of $\mathcal{A}$ and C^* and substituting D_1^* for D_1. Variation with respect to D_1 yields no new condition. These conditions leave much freedom. [11]

But practicality requires that one imposes restrictions of the form $\mathcal{A}(z,x) = 0$ except when z is sufficiently close to x, and similarly for C and D_1. (We will regard Λ^1 as a sublattice of Λ^0.) The proposal is to find the optimal kernels which satisfy these constraints by minimizing E_1.

---

[10] For instance, one may admit link variables $U(z,w)$ in the covariant Laplacian Δ which take values 0 or 1. Links (z,w) with $U(z,w) = 0$ are considered deleted from the lattice. If deleted links are distributed at random with suitable probability, on obtains a fractal lattice. Solutions of the equation $(-\Delta + M^2)\chi_0 = f_0$ admit a random walk representation on this lattice.

[11] For instance, one may demand that $C = (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}^*$ or, equivalently, $\mathcal{A} = C^*(CC^*)^{-1}$. Then $C\mathcal{A} = \mathbb{1}$, the second condition follows from the first, and Γ agrees with the fluctuation field propagator (40) with $D_1 = -CC^*\Delta_{eff}$ if $D_0 = -\Delta$.

Figure 9: Neural net from a twogrid (Λ^0, Λ^1) in one dimension. It has one input layer Λ^0, one output layer Λ^0 and one layer Λ^1 of hidden nodes.

The idea was tested on a baby model, the 1-dimensional Poisson equation

$$(-\Delta + M^2)\ \chi_0 = f_0 \tag{107}$$

on a chain Λ^0 of lattice spacing 1, with very small M. Λ^1 consisted of the even sites in Λ^0. As a condition of practicality it was demanded that

$$C(x,z) = 0 = \mathcal{A}(z,x) \quad \text{unless} \quad |z-x| \leq 1 \ .$$

Optimization via gradient descent yielded the expected result, in the limit of small M,

$$C(x,x) = 1 \quad , \quad C(x, x \pm 1) = \frac{1}{2} \quad , \quad D_1 = 2(-\Delta_1 + 4M^2) \tag{108}$$

$$\mathcal{A}(z,x) = C(x,z) \qquad (x \in \Lambda^1, z \in \Lambda^0) \ , \tag{109}$$

modulo normalization conventions. Δ_1 stands for the Laplacian on Λ^1.

Neural net approach

It is interesting to rephrase the problem in the language of neural nets [41].

We regard the layers of the multigrid as layers of a perceptron. The sites of the lattices are called nodes in this context. In the twogrid case there will be an input layer with nodes $z \in \Lambda^0$, a hidden layer with nodes $x \in \Lambda^1$ which are sites of the coarse grid, and an output layer with nodes $w \in \Lambda^0$. Nodes z of the input layer are connected to nodes x of the hidden layer by "neurons" with connection strength $C(x,z)$, so that node x receives as input a linear superposition of the output of the input nodes z. Nodes x of the hidden layer are connected to nodes w of the output layer by "neurons" of connection strength $\mathcal{A}(w,x)$ (see Fig. 9).

The network receives as input an approximation $\xi(z)$ to the solution of the equation (94) which we wish to solve, and performs an iteration step in order to obtain as output an improved approximation $O(z)$. The desired output is the exact solution $\zeta(z) = D_0^{-1} f_0$. If the iteration step involves the exact solution of the coarse grid equation (98) then the network is not strictly feed-forward. [12] This is because the hidden layer needs to do computations more general than the usual computation of some nonlinear function at individual nodes in order to solve the

[12] A neural network is called "feed-forward" if the nodes in a layer receive their only input from nodes on preceding layers [41].

coarse grid equation. This could be remedied if one envisages iterative solution of the coarse grid equation (98). We return to this below.

The network is supposed to learn to compute an output O which is as close as possible to the target ζ. The learning process consists in adjusting the connection strengths $C(x,z)$ and $\mathcal{A}(z,x)$. One envisages that a training set (ξ^μ, ζ^μ) is presented to the network to teach it. It involves a sequence of inputs ξ^μ. The network has its output O^μ compared to the desired answer ζ^μ and receives feedback about the error which is then used to adjust the connection strengths according to some learning rule. Application of typical learning rules amounts to iterative minimization of some cost functional E. The standard cost functional is

$$E = \sum_\mu \sum_z |O^\mu(z) - \zeta^\mu(z)|^2 \ .$$

In our case, the output is a linear function of the input and of f_0 which has the exact solution $D_0^{-1} f_0$ as a fixed point. Therefore

$$O = \rho\ \xi + (1-\rho) D_0^{-1} f_0 \ ,$$

where ρ is the iteration matrix. The target $\zeta^\mu = D_0^{-1} f_0$ is independent of the input. The cost functional becomes

$$E = \sum_\mu \left\| \rho \left(\xi - D_0^{-1} f_0 \right) \right\|^2$$

where $\|f\|^2 = \langle f, f\rangle = \sum_z |f(z)|^2$. Since the equation is linear, we may imagine without loss of generality that f_0 is arbitrarily small. Assuming the training input is a complete orthonormal set ξ^μ of functions on Λ^0 in the limit of small f_0, the cost functional becomes

$$E = \|\rho\|^2 \equiv \mathrm{Tr}(\rho\rho^*) \ .$$

This is the same cost functional as considered before.

We could restrict our attention to the coarse grid updating whose updating matrix is ρ_1. In this case one wants mainly to suppress the low frequency modes of the error. This motivates us to propose a training set with ξ^μ such that $D_0\xi^\mu$ are orthonormal. As a result one finds that E_1 should be minimized, where $E_1 = \|D_0^{-1}\rho_1\|^2$ as before. Minimization of E_1 leads to less practical learning rules than E because of the presence of the operator D_0^{-1}, though.

Let us return to the question of the implementation within the framework of feed-forward networks.
We assume that the input layer knows to compute $r_0(z) = D_0\xi(z) - f_0(z)$ from the input ξ. This is easily implemented within the standard feed-forward framework by adding an extra layer. Initial fine grid relaxation sweeps to smoothen r_0 are also easily implemented. The nodes x of the hidden layer will receive input $r_1(z) = \sum_z C(x,z) r_0(z)$ that is computed from the smoothened version of r_0.

Let us imagine that we do not use the exact solution e_1 to the coarse grid equation (98), but content ourselves with an approximation $\Xi(x)$ to $-e_1(x)$ which comes from a single Jacobi relaxation sweep. Then the node x is asked to compute the output

$$\Xi(x) = -d_1(x)^{-1}\ r_1(x) \ , \qquad (110)$$

¿from its input $r_1(x)$, which it can legally do. d_1 is the diagonal part of D_1, viz.

$$D_1 f(x) = d_1(x) f(x) + \text{ contributions } \propto f(y) \text{ with } y \neq x \ .$$

We imagine that there are additional connections (not shown in figure (9)) of strength 1 from input nodes to output nodes which represent the same lattice point z. The total input to the output node z is then

$$O(z) = \xi(z) + \sum_x \mathcal{A}(z,x)\ \Xi(x) \approx \xi(z) - \sum_x \mathcal{A}(z,x)\ e_1(x)\ , \tag{111}$$

and this is also the output from the output node z.

Updating of connection strengths $C(x,z)$ and $\mathcal{A}(x,z)$ proportional to the gradient of the cost functional E leads to a standard *back-propagation learning rule* as described in [41]. As is discussed there (Sect. 6.1) the same network - or rather a bidirectional version of it - can be used to compute the necessary adjustments.

There is still one nonstandard feature. The network should also learn $d_1(x)^{-1}$. This becomes a standard learning problem (to be solved by the back-propagation learning rule) if we think of duplicating the hidden layer, connecting the duplicates of site x by "neurons" of connection strength $d_1(x)^{-1}$.

We discussed a twogrid for simplicity. Generalization to a multigrid is obvious and will be presented elsewhere [42].

10 Blocked Gauge Fields

Remember that the gauge covariant Laplace operator Δ for gauge field U has kernel ($a = 1$)

$$\Delta(z,w) = \begin{cases} -2d & \text{if } z = w, \\ U(z,w) & \text{if } z \text{ n.n. } w\ . \end{cases} \tag{112}$$

in d dimensions. And similarly for the staggered Dirac operator[13]

$$\not\nabla(z,w) = \begin{cases} \eta_\mu(z)\, U(z,w) & \text{if } w = z + \frac{1}{2}e_\mu,\ \mu = -d,\dots,d, \mu \neq 0, \\ 0 & \text{otherwise}\ . \end{cases} \tag{113}$$

A blockspin definition for matter fields gives us the effective operators on coarser layers, $\Delta_{\text{eff}} = \mathcal{A}^*\Delta\mathcal{A}$ and $\not\nabla_{\text{eff}} = \mathcal{A}^*\not\nabla\mathcal{A}$. Given Δ_{eff} or $\not\nabla_{\text{eff}}$, we have candidates for gauge fields $\mathcal{U}$ on the block lattice,

$$\begin{aligned} \mathcal{U}(x,y) = &\ \ \Delta_{\text{eff}}(x,y) \quad \text{for } x,y \text{ nearest neighbours} \\ \text{resp.}\quad \mathcal{U}(x,y) = &\ \ \not\nabla_{\text{eff}}(x,y)\,/\,\eta_\mu(x) \quad \text{if } y = x + L_b\tfrac{1}{2}e_\mu\ . \end{aligned} \tag{114}$$

These expressions for $\mathcal{U}(x,y)$ have the correct gauge covariance property.

We add some comments on the Dirac case. Since we use odd L_b, the staggered block lattice can be regarded as a sublattice of Λ in which the pseudoflavor of a site does not depend on whether we regard it as a site of the fundamental or the block lattice, and the same is true of its property of being even or odd. The kernel $\mathcal{A}$ will always have the property that $\mathcal{A}(z,x) = 0$ unless z and x are both even or both odd. Therefore $\not\nabla_{eff}$ connects even to odd sites, The factor $\eta_\mu(x)^{-1}$ can be regarded as defined by regarding x as a point in a sublattice of the fundamental lattice. This agrees with the natural definition on the block lattice. Note also that $\eta_\mu(z)$ is independent of the μ-component z_μ) of z.

[13]We use the conventions $\eta_{-\mu} = -\eta_\mu$ and $e_{-\mu} = -e_\mu$. z is called odd if the sum of its integer coordinates is odd.

The blocked gauge fields $\mathcal{U}(x,y)$ are actually "dielectric" gauge fields [19]. They may be brought back into the gauge group by multiplying with a positive real number (for $U(1)$ and $SU(2)$) or with a positive matrix (for $U(N), N>2$).

A different choice of blockspin for gauge fields was introduced by Balaban and used in rigorous renormalization group work [43].

11 Simulation Methods for the Computation of Effective Actions

The standard Monte Carlo simulation techniques for the computation of functional integrals can be used to compute effective actions, and in particular effective potentials.

There are three methods to take into account the δ-function constraints.

1.) **Gaussian blockspins.** Use a Gaussian in place of the δ-function in the constraint. This definition of a block spin transformation is very convenient for numerical work. But it has some inherent dangers which have been discussed in the literature (cp. [44]): Fictitious marginal operators may appear and cloud the renormalization group flow. Some numerical results are found in [44].

2.) **Simulations at fixed blockspin Φ** [44]. Do the updating so that Φ is fixed. Compute

$$\langle \partial\mathcal{H}/\partial\Phi(x)\rangle_\Phi$$

This determines $\mathcal{H}_{eff}(\Phi)$ up to a constant.

For linear blockspin transformations, the quantity $\partial\mathcal{H}/\partial\Phi(x)$ is defined as follows:

$$\frac{\partial\mathcal{H}}{\partial\Phi(x)} = \int_z \frac{\partial\mathcal{H}(\varphi)}{\partial\varphi(z)}\mathcal{A}(z,x) \tag{115}$$

so that the action $\mathcal{H}$ changes by $\delta\mathcal{H} = \frac{\partial\mathcal{H}}{\partial\Phi(x)}\delta\Phi(x)$ when φ is changed by an amount $\delta\varphi(z) = \mathcal{A}(z,x)\delta\Phi(x)$. Simulations with fixed block spin are simulations for an auxiliary theory with an infrared cutoff a_0^{-1}. Therefore the critical slowing down will be governed by the ratio a_0/a which is usually appreciably less than the correlation length in units of a.

This method is general, but it is costly when one wants to compute more than effective potentials.

3.) **Method of "fluctuating coupling constants"** [9, 15] This method is applicable in cases such as Higgs models with and without gauge fields, where the effective action depends only on matter fields Φ which are related by a linear blockspin definition to matter fields which appear *polynomially* in $\mathcal{H}$.

Given φ, U then the high frequency field ζ is defined. A sequence of configurations (φ_a, U_a) yields therefore a sequence (ζ_a, U_a), $a = 1, 2, \ldots, N_{conf}$. Expanding

$$\mathcal{H}(U_a, \zeta_a + \mathcal{A}\Phi) = \sum_m g_a^m \Phi^m \tag{116}$$

defines a sequence g_a^m of "fluctuating coupling constants". They are stored. At the end of the run, the effective potential can be recovered from them. For a 1-component field

$$e^{-\mathcal{H}_{eff}(\Phi)} = N_{conf}^{-1}\sum_a Z_a^{-1}\exp\left[-\sum_m g_a^m \Phi^m\right] \tag{117}$$

where Z_a is the 1-dimensional integral

$$Z_a = \int d\Phi \exp\left[-\sum_m g_a^m \Phi^m\right] \tag{118}$$

One may choose to expand around another point than $\Phi = 0$ instead.

In principle, fluctuation coupling constants could also be used to encode the information about full effective actions. But in this case g_a^m would be functions of m arguments $x_1, \ldots, x_m$ on the block lattice. Only few of them could possibly be stored.

Multigrid Updating of Matter Fields

The multigrid method provides a tool to fight critical slowing down in Monte Carlo simulations. It uses or is motivated by constructs that we discussed. Therefore we would like to mention it at least briefly.

Imagine that the blockspin Φ^j attached to some block lattice Λ^j of lattice spacing a_j is changed at site $x \in \Lambda^j$ by $\delta\Phi^j(x)$. This will induce a change of the matter field

$$\delta\varphi(z) = \mathcal{A}^j(z,x)\delta\Phi^j(x) \ .$$

$\mathcal{A}^j$ should be thought of as minimizing the kinetic energy subject to some constraint (such as $C\varphi = \Phi$), therefore it should be *smooth in* z (no jumps). If we do not specify the averaging kernel C, there are no further restrictions on $\mathcal{A}$.

Similar updatings are possible for fields subject to constraints such as unit vectors in σ-models, spins in CP^n-models etc. For instance

$$\varphi(z) \mapsto \exp\left[\mathcal{A}(z,x)\ \epsilon\right]\ \varphi(z) \ , \tag{119}$$

where the matrix ϵ is an element of the Lie algebra of the symmetry of the model.

In the "unigrid" approach of Hasenbusch, Mack and Meyer [45], hypercubes x of different side lengths are considered and updatings of the form (119) are performed with smooth kernels $\mathcal{A}(z,x)$ which vanish for z outside x. The positions of the hypercubes x are chosen at random in the lattice. It turns out to be important for the effectiveness of the method that the selected hypercubes x of one length have some overlap.

The unigrid method proved very successful for 2-dimensional asymptotically free models: The critical exponent z which governs critical slowing down is nearly zero in this case. References are [46] for the CP^3-model and [45] for the $O(3)-$model. It is essential for the success that these models have no mass parameter [47]. The effective potential for 4-dimensional φ^4-theory was computed using a similar updating scheme [15], but with disjoint hypercubes x which form a lattice, and with kernels $\mathcal{A}$ which had support on hypercubes x and on their nearest neighbours.

Acknowledgements

Support by Deutsche Forschungsgemeinschaft is gratefully acknowledged. For providing resources and advice, we are indebted to HLRZ/Jülich and its staff. G.M. and M.S. thank the organizers of the 31.IUKT in Schladming for their efforts in creating the pleasant and stimulating atmosphere during the school. One of us (M.S.) thanks Thorsten Ohl for gnuidance.

Appendix: Constraint Effective Potential with Gauge Fields in Finite Volume

The result of the 1-loop calculation of Sect. 6 reads as follows (for details see [48])

$$U_V(\Phi) = U_V^{(0)}(\Phi) + \frac{1}{2V}\,\mathrm{Tr}\,\left[\ln\left\{\frac{1}{2}\,\frac{\delta}{\delta A_{rz\mu}}\,\frac{\delta}{\delta A_{r\bar{z}\nu}}\left(S_g^{(q)}[\Phi,A]\right)\right\}\right]_{A=0} \tag{120}$$

with

$$U_V^{(0)}(\Phi) = U_{cl}(\Phi) + \frac{1}{2V}{\sum_n}' \ln\left(\widehat{k}_n^2 + U''_{cl}(\Phi)\right) + \frac{3}{2V}{\sum_n}' \ln(\widehat{k}_n^2 + U'_{cl}(\Phi)/\,\Phi) \tag{121}$$

$$S_g^{(q)}[\Phi,A] = S_g^q[A] + U_V^{(2)}[\Phi,A] \tag{122}$$

where

$$\begin{aligned} U_V^{(2)}(\Phi,A) = \quad & \tfrac{1}{4V}(U'_{cl}(\Phi)/\Phi){\sum_n}'\tfrac{1}{(\widehat{k}_n^2)^2}\sum_r |A_{rno}|^2 + \tfrac{1}{4V}{\sum_n}'(\Gamma_{n\phi_0}(\Phi) + 3\Gamma_{n\phi_r}(\Phi)) \\ & \sum_r (A^2_{rnn} + \tfrac{1}{\widehat{k}_n^2}|A_{rno}|^2) - \tfrac{1}{4V^2}{\sum_{n,m}}'(2\Gamma_{n\phi_r}(\Phi)\Gamma_{m\phi_r}(\Phi) + \Gamma_{n\phi_0}(\Phi)\Gamma_{m\phi_r}(\Phi) \\ & +\Gamma_{n\phi_r}(\Phi)\Gamma_{m\phi_0}(\Phi))\sum_r |A_{rnm}|^2 + \tfrac{\lambda_0}{12}\Phi^2{\sum_n}'(4\Gamma_{n\phi_r}(\Phi) + \Gamma_{n\phi_0}(\Phi)) \\ & \tfrac{1}{(\widehat{k}_n^2)^2}\sum_r |A_{rno}|^2 \end{aligned} \tag{123}$$

with

$$\begin{cases} U_{cl}(\Phi) = \frac{1}{2}\,m_0^2\,\Phi^2 + \frac{\lambda_0}{4!}\Phi^4, \quad \widehat{k}_n^2 = 4\sum_\mu \sin^2(\frac{\pi}{L}n_\mu), \quad \frac{L}{2} < n_\mu \le \frac{L}{2} \\ \Gamma_{n\varphi_0}(\Phi) = (\widehat{k}_n^2 + U''_{cl}(\Phi))^{-1}, \qquad \Gamma_{n\varphi_r}(\Phi) = (\widehat{k}_n^2 + U'_{cl}(\Phi)/\Phi)^{-1} \end{cases} \tag{124}$$

Φ the absolute value of the covariant magnetization of (62) and $A_{rno}, A^2_{rnn}, A_{rnm}$ defined by

$$\begin{cases} (A^2_r)_{oo} & = \frac{1}{V}\sum_{z\mu}\psi_0^*(z)A_{rz-\mu\mu}A_{rz-\mu\mu}\psi_0(z-\mu) \\ (A^2_r)_{no} & = \frac{1}{V}\sum_{z\mu}\psi_n^*(z)A_{rz-\mu\mu}A_{rz-\mu\mu}\psi_0(z-\mu) \\ A_{rnm} & = \frac{1}{V}\sum_{z\mu}\psi_n^*(z)A_{rz-\mu\mu}\,\psi_m\,(z-\mu) \\ |A_{rno}|^2 & = A_{rno}A^*_{rno} \end{cases} \tag{125}$$

Here, $\psi_n(z) = \exp(\frac{2\varphi i}{L}nz)$ and the eigenvalue ϵ_0 is given by (70). U'_{cl} and U''_{cl} are the first and second derivative of U_{cl}, and prime $'$ on the sum means that the $n=0$ term is omitted.

The trace in (120) involves summation over the discrete $O(4)$-matrix indices and a sum over space-time points on a lattice Λ.

References

[1] M. Göpfert and G. Mack, Commun. Math. Phys. 82 (1982) 545

[2] G. Mack, M. Pordt, Commun. Math. Phys. 97 (1985) 267

[3] M. Karliner, G. Mack, Nucl. Phys. B225 [FS9] (1983) 371

[4] K. Gawędzki and A. Kupiainen, Commun. Math. Phys. 77 (1980) 31

[5] H.B. Nielsen, M. Ninomiya, Phys. Lett. B105 (1981) 219

[6] L. O'Raifeartaigh, A. Wipf and H. Yoneyama, Nucl. Phys. B271 (1986) 653

[7] K. Wilson, Phys. Rev. D3 (1971) 1818; Rev. Mod. Phys. 55 (1983) 583;
J. Kogut and K. Wilson, Phys. Rep. 12C (1974) 75

[8] J. Polchinski, Nucl. Phys. B231 (1984) 269

[9] G. Mack, Multigrid Methods in Quantum Field Theory , Cargèse Lectures July 1987, *in:* G. 't Hooft et al., Nonperturbative Quantum Field Theory (Plenum Press, New York, 1989)

[10] K. Gawędzki and A. Kupiainen, Commun. Math. Phys. 99 (1985) 197

[11] A. Ringwald, C. Wetterich, Nucl. Phys. **B 334** (1990) 506;
C. Wetterich, Nucl. Phys. **352** (1991) 529;
C. Wetterich, The average action for scalar fields near phase transitions, DESY 91-088 (August 91)

[12] G. Palma, Finite Size Scaling Analysis of the Constraint Effective Potential computed from Multigrid Monte Carlo, DESY 90-083 (July 1990)

[13] G. Palma, Renormalized Loop Expansion to Compute Finite Size Effects of the Constraint Effective Potential, to appear in Z. Phys.

[14] G. Palma, Renormalized Loop Expansion for the Gauge Covariant Constraint Effective Higgs Potential on the Lattice, PhD-thesis Hamburg, DESY 92–031 (February 1992)

[15] G. Mack and S. Meyer, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 293

[16] L. Lin, J. Kuti and Y. Shen, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 26

[17] M. Lüscher and P. Weisz, Nucl. Phys. B290 [FS20] (1987) 25;
M. Lüscher and P. Weisz, Nucl. Phys. B295 [FS21] (1988) 65

[18] T. Kalkreuter, Projective block spin transformations in lattice gauge theory, preprint, DESY 90–158 (December 1990) to appear in Nucl. Phys. B

[19] G. Mack, Nucl. Phys. B235 [FS11] (1984) 197

[20] C. Wetterich and M. Reuter, Average Action for the Higgs Model with Abelian Gauge Symmetry, preprint DESY 92–037 (March 1992)

[21] J. Klauder, lectures at this school

[22] I. Montvay, Nucl. Phys. B293 (1987) 479

[23] A. Hasenfratz and P. Hasenfratz, Phys. Rev. D34 (1986) 3160

[24] S. Coleman and E. Weinberg, Phys. Rev. D7 (1973) 1888

[25] T. DeGrand, lectures at this school

[26] T. Kalkreuter, G. Mack and M. Speh, Int. J. Mod. Phys. C3 (1992) 121

[27] G.Mai, Ein Blockspin für $2^{d/2}$ Fermionen, diploma thesis, Hamburg (October 1989)

[28] H. Joos and M. Schaefer, Z. Phys. C34 (1987) 465;
M.F.L. Golterman, Nucl. Phys. B278 (1986) 417

[29] J. Smit, Acta Phys. Polon. B 17 (1986) 531

[30] R. Ben-Av, A. Brandt and S. Solomon, Nucl. Phys. B329 (1990) 193;
R. Ben-Av, A. Brandt, M. Harmatz, E. Katznelson, P.G. Lauwers, S. Solomon and K. Wolowesky, Phys. Lett. B253 (1991) 185; Nucl. Phys. B (Proc. Suppl.) 20 (1991) 102;
R. Ben-Av, P.G. Lauwers and S. Solomon, Nucl. Phys. B 374 (1992) 249;
P.G. Lauwers and S. Solomon, Int. J. Mod. Phys. C3 (1992) 149

[31] W. Hackbusch, Multi-grid Methods and Applications (Springer-Verlag, Berlin, 1985)

[32] A. Hulsebos, J. Smit and J.C. Vink, Int. J. Mod. Phys. C3 (1992) 161
J.C. Vink, Phys. Lett. B272 (1991) 81; Nucl. Phys. B (Proc. Suppl.), to appear (LATTICE '91)

[33] R.C. Brower, K.J.M. Moriarty, E. Myers and C. Rebbi, in: Multigrid Methods, ed. S.F. McCormick (Marcel Dekker, New York, 1988)

[34] R.C. Brower, C. Rebbi and E. Vicari, Phys. Rev. D43 (1991) 1965; Phys. Rev. Lett. 66 (1991) 1263;
R.C. Brower, K.J.M. Moriarty, C. Rebbi and E. Vicari, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 89; Phys. Rev. D43 (1991) 1974

[35] A. Hulsebos, J. Smit and J.C. Vink, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 94; Nucl. Phys. B368 (1992) 379

[36] R.C. Brower, R.G. Edwards, C. Rebbi and E. Vicari, Nucl. Phys. B366 (1991) 689

[37] V. Vyas, Int. J. Mod. Phys. C3 (1992) 169

[38] T. Kalkreuter, Phys. Lett. B276 (1992) 485

[39] A. Brandt, Multigrid techniques: 1984 guide, GMD-Studie No. 85

[40] T. Kalkreuter, *in preparation*

[41] J. Hertz, A. Krogh and R.G. Palmer, Introduction to the Theory of Neural Computation (Addison-Wesley, Redwood City CA, 1991)

[42] G. Mack and M. Speh, *in preparation*

[43] T. Balaban, Commun. Math. Phys. 89 (1983) 571

[44] K. Pinn, Z. Phys. C45 (1990) 453;

K. Pinn, Z. Phys. C47 (1990) 325;

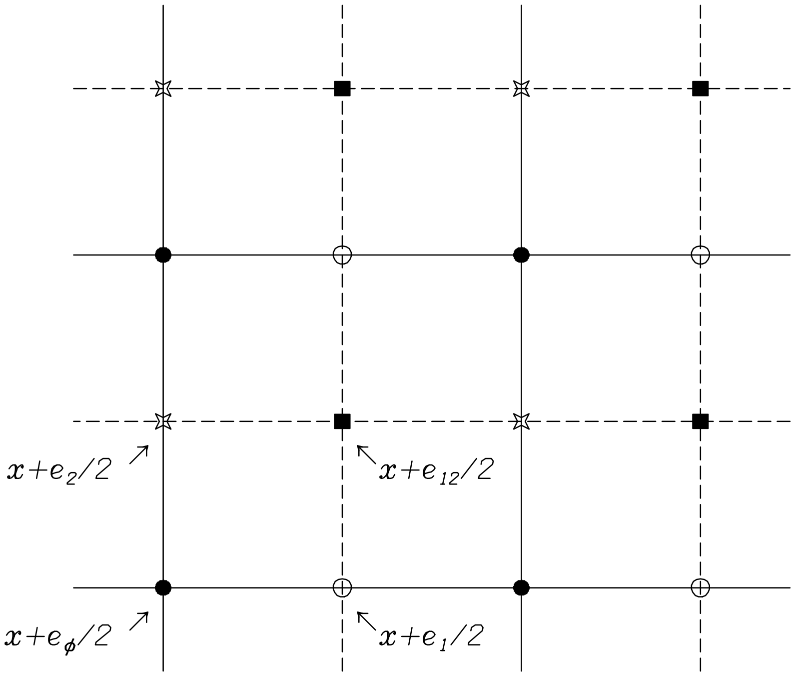
Computation of Effective Hamiltonians by Monte Carlo Simulations with Fixed Blockspins, preprint DESY 88–123 (August 1988)

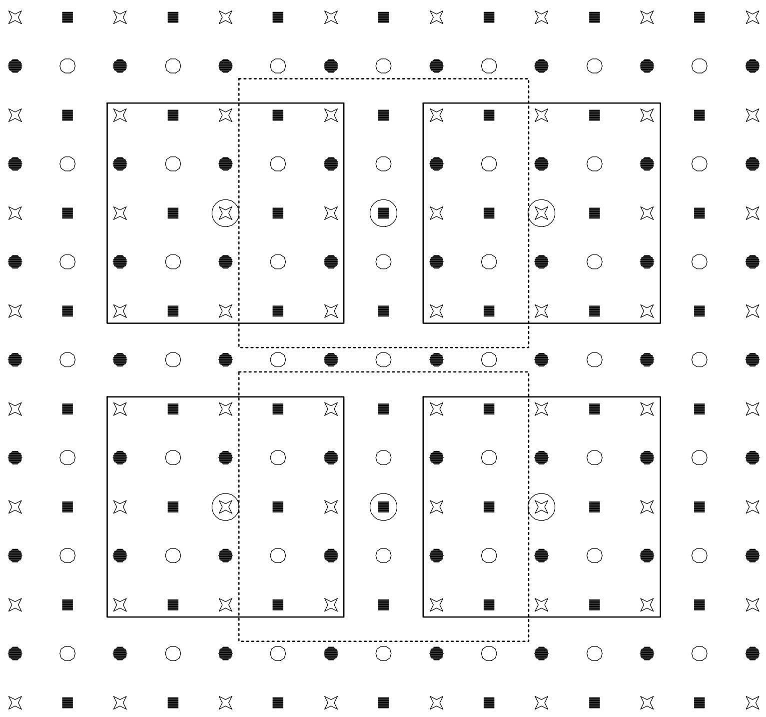
[45] M. Hasenbusch, G. Mack and S. Meyer, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 110

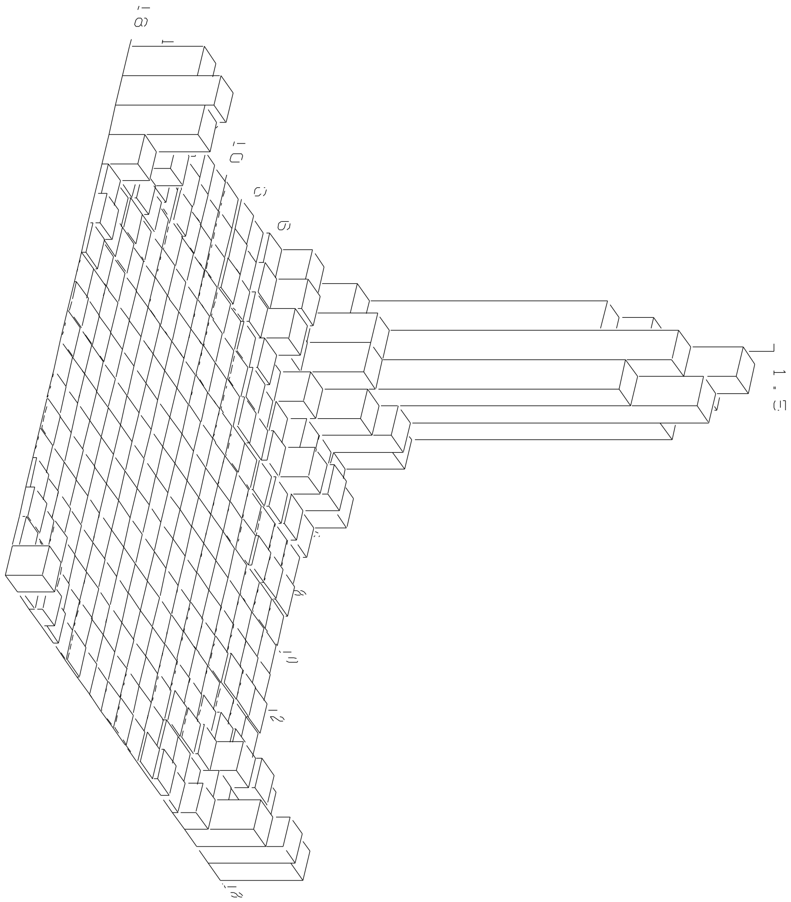
[46] S. Meyer, M. Hasenbusch, Testing accelerated algorithms in the lattice CP^3 model, preprint Kaiserslautern-TH-91-23 (January 1992)

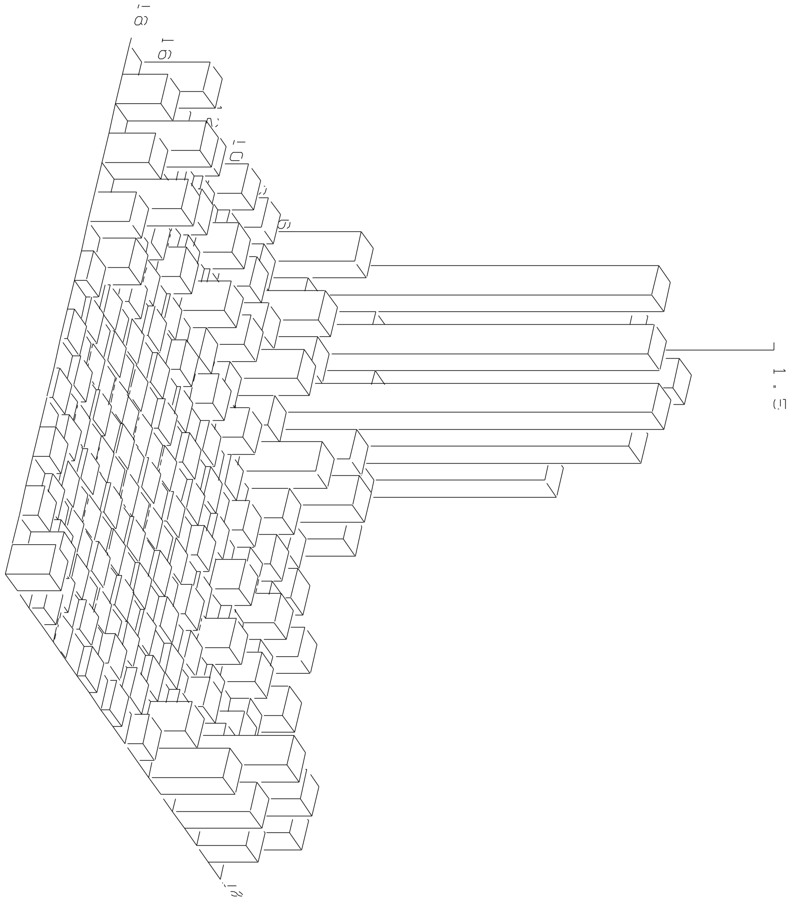
[47] M. Grabenstein and K. Pinn, Acceptance Rates in Multigrid Monte Carlo, preprint DESY 91–157 (December 1991), to appear in Phys. Rev. D

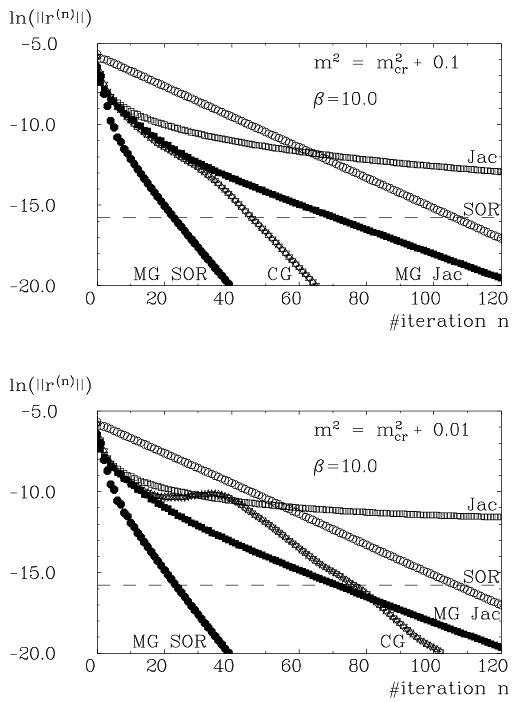
[48] G. Palma, *in preparation.*


$x+e_2/2$
$x+e_{12}/2$
$x+e_\phi/2$
$x+e_1/2$



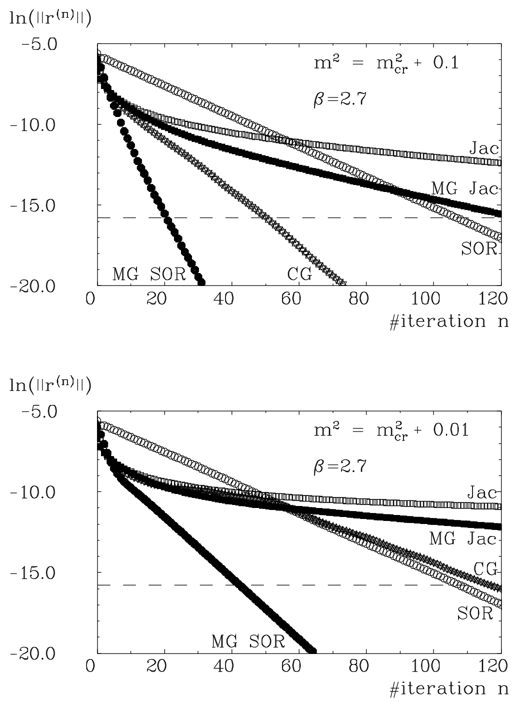






ln(||r(n)||)
-5.0
-10.0
-15.0
-20.0
0
20
40
60
80
100
120
#iteration n
m² = m²cr + 0.1
β=10.0
Jac
SOR
MG SOR
CG
MG Jac
ln(||r(n)||)
-5.0
-10.0
-15.0
-20.0
0
20
40
60
80
100
120
#iteration n
m² = m²cr + 0.01
β=10.0
Jac
SOR
MG Jac
MG SOR
CG

Fig. 1 (a)



$\ln(\|r^{(n)}\|)$
$m^2 = m_{cr}^2 + 0.1$
β=2.7
Jac
MG Jac
SOR
MG SOR
CG
-5.0
-10.0
-15.0
-20.0
0
20
40
60
80
100
120
#iteration n
$\ln(\|r^{(n)}\|)$
$m^2 = m_{cr}^2 + 0.01$
β=2.7
Jac
MG Jac
CG
SOR
MG SOR
-5.0
-10.0
-15.0
-20.0
0
20
40
60
80
100
120
#iteration n

Fig. 1 (b)

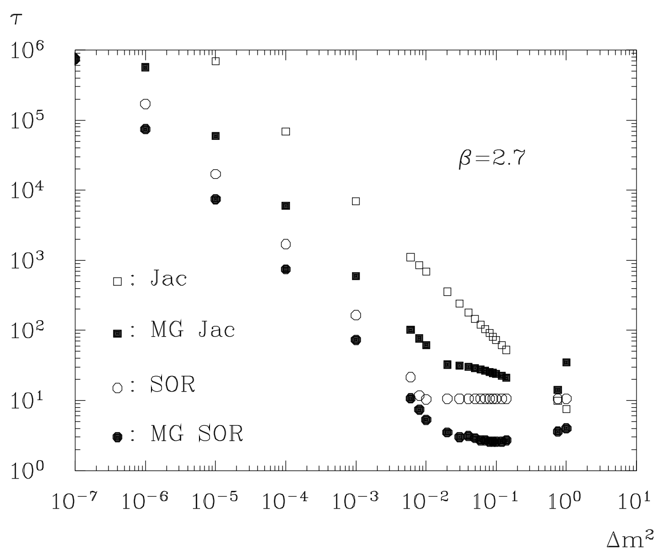
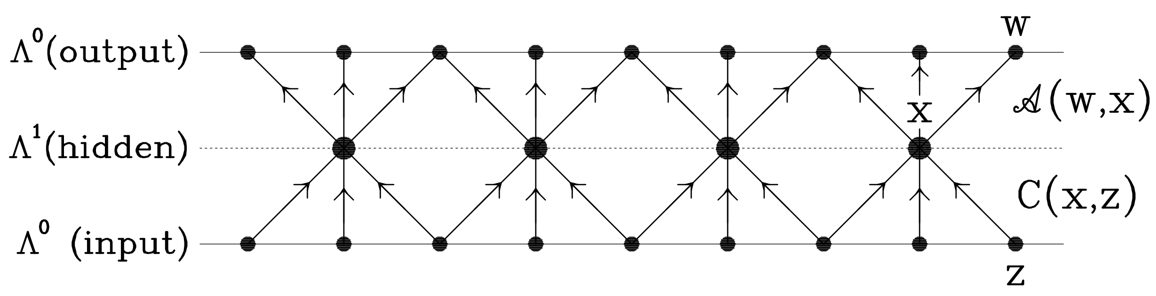

τ
β=2.7
□ : Jac
■ : MG Jac
○ : SOR
● : MG SOR
10^0
10^1
10^2
10^3
10^4
10^5
10^6
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^0
10^1
Δm^2

Fig. 2



Λ^0(output)
Λ^1(hidden)
Λ^0 (input)
w
x
z
$\mathscr{A}(w,x)$
$C(x,z)$