An introduction to Symanzik's O(a) improvement programme

Stefan Sint

Trinity College Dublin

STRONGnet Summer School 2011

Bielefeld, 14 June 2011

Contents

Non-perturbative definition of QCD (1)

To define QCD as a QFT it is not enough to write down its classical Lagrangian:

$$\mathcal{L}_{\rm QCD}(x) = \frac{1}{2g^2}\,\mathrm{tr}\,\{F_{\mu\nu}(x)F_{\mu\nu}(x)\} + \sum_{i=1}^{N_{\rm f}} \overline{\psi}_i(x)\,(\not{D} + m_i)\,\psi_i(x)$$

One needs to define the functional integral:

- Introduce a Euclidean space-time lattice and discretise the continuum action such that the doubling problem is solved
- Consider a finite space-time volume $\Rightarrow$ the functional integral becomes a finite dimensional ordinary or Grassmann integral, i.e. mathematically well defined!
- Take the infinite volume limit $L \rightarrow \infty$
- Take the continuum limit $a \rightarrow 0$

Non-perturbative definition of QCD (2)

- In massive theories the infinite volume limit is reached with exponential corrections $\Rightarrow$ not a major problem in practice.
- Continuum limit: existence only established order by order in perturbation theory & only for selected lattice regularisations:
  - lattice QCD with Wilson quarks [Reisz '89]
  - lattice QCD with overlap/Neuberger quarks [Reisz, Rothe '99]
  - not (yet ?) for lattice QCD with staggered quarks [cf. Giedt '06]
- From asymptotic freedom expect

$$g_0^2 = g_0^2(a) \quad \stackrel{a\rightarrow 0}{\sim} \quad \frac{-1}{2b_0 \ln a}, \qquad b_0 = \frac{11N}{3} - \frac{2}{3}N_\mathrm{f}$$

Non-perturbative definition of QCD (3)

Working hypothesis: the perturbative picture is essentially correct:

- The continuum limit of lattice QCD exists and is obtained by taking $g_0 \rightarrow 0$
- Hence, QCD is asymptotically free, naive dimensional analysis applies: non-perturbative renormalisation of QCD is based on the very same counterterm structure as in perturbation theory!
- Absence of analytical methods: try to take the continuum limit numerically, i.e. by numerical simulations of lattice QCD at decreasing values of g_0.
- Typical current ranges for a:

$$a = 0.04 - 0.1 \text{ fm}$$

  i.e. at most a variation by a factor 2-3!

Renormalisation of QCD

- The basic parameters of QCD are g_0 and m_i, $i = u, d, s, c, b, t$.
- To renormalise QCD one must impose a corresponding number of renormalisation conditions
- We only consider gauge invariant observables $\Rightarrow$ no need to consider field renormalisations for quark, gluon or ghost fields or the renormalisation of the gauge parameter.
- All physical information (particle masses and energies, particle interactions) is contained in the (Euclidean) correlation functions of gauge invariant composite, local fields $\phi_i(x)$

$$\langle \phi_1(x_1) \cdots \phi_n(x_n) \rangle$$

- a priori each field ϕ_i requires renormalisation, and thus further renormalisation conditions.

Counterterm structure in lattice QCD with Wilson quarks

The action $S = S_{\rm f} + S_{\rm g}$ is given by

$$S_{\rm f} = a^4 \sum_x \overline{\psi}(x) \left(D_W + m_0\right) \psi(x), \qquad S_{\rm g} = \frac{1}{g_0^2} \sum_{\mu,\nu} \mathrm{tr}\, \{1 - P_{\mu\nu}(x)\}$$

$$D_W = \tfrac{1}{2} \left\{ \left(\nabla_\mu + \nabla_\mu^*\right) \gamma_\mu - a \nabla_\mu^* \nabla_\mu \right\}$$

- Symmetries: $\mathrm{U}(N_{\rm f})_{\rm V}$ (mass degenerate quarks), P, C, T and $O(4, \mathbb{Z})$

$\Rightarrow$ Renormalized parameters:

$$g_{\rm R}^2 = Z_g g_0^2, \qquad m_{\rm R} = Z_m \left(m_0 - m_{\rm cr}\right), \qquad am_{\rm cr} = am_{\rm cr}(g_0).$$

- In general: $Z = Z(g_0, a\mu, am_0)$;
- Quark mass independent renormalisation schemes: $Z = Z(g_0, a\mu)$
- Simple non-singlet composite fields, e.g. $P^a = \overline{\psi}\gamma_5 \frac{1}{2}\tau^a \psi$ renormalise multiplicatively, $P_{\rm R}^a = Z_{\rm P}(g_0, a\mu, am_0) P^a$

Symanzik's effective continuum theory (1) [Symanzik '79]

- goal: render the a-dependence of lattice correlation functions explicit. $\Rightarrow$ structural insight into the nature of cutoff effects
- at scales far below the cutoff a^{-1}, the lattice theory is effectively continuum-like and can be represented by an effective continuum theory, with action

$$
\begin{aligned}
S_{\rm eff} &= S_0 + aS_1 + a^2 S_2 + \ldots, \qquad S_0 = S_{\rm QCD}^{\rm cont} \\
S_k &= \int \mathrm{d}^4 x\, \mathcal{L}_{\rm k}(x)
\end{aligned}
$$

$\mathcal{L}_{\rm k}(x)$: linear combination of fields

- with canonical dimension $4+k$
- which share all the symmetries with the lattice action

Symanzik's effective continuum theory (2)

The first term $\mathcal{L}_1$ can be parametrized by:

$$\mathcal{L}_1 = c_1^{(1)}\overline{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi + c_1^{(2)}\overline{\psi}D^2\psi + c_1^{(3)}m\,\overline{\psi}\not{D}\psi + c_1^{(4)}m^2\overline{\psi}\psi + c_1^{(5)}m\,\mathrm{tr}\,\{F_{\mu\nu}F_{\mu\nu}\}$$

The same procedure applies to composite fields:

$$\phi_{\rm eff}(x) = \phi_0 + a\phi_1 + a^2\phi_2 \ldots$$

for instance: $\phi(x) = P^a(x)$:

$$P^a_{\rm eff} = \overline{\psi}\gamma_5\tfrac{1}{2}\tau^a\psi + c_P^{(1)}m\,\overline{\psi}\gamma_5\tfrac{1}{2}\tau^a\psi + c_P^{(2)}\left(\overline{\psi}\overleftarrow{\not{D}}\gamma_5\tfrac{1}{2}\tau^a\psi - \overline{\psi}\gamma_5\tfrac{1}{2}\tau^a\not{D}\psi\right)$$

Consider renormalised, connected lattice n-point functions of a multiplicatively renormalisable field ϕ

$$G_n(x_1,\ldots,x_n) = Z_\phi^n\langle\phi(x_1)\cdots\phi(x_n)\rangle_{\rm con}^{lat}$$

Symanzik's effective continuum theory (3)

Effective field theory description:

$$\begin{aligned} G_n(x_1,\ldots,x_n) &= \langle\phi_0(x_1)\ldots\phi_0(x_n)\rangle_{\rm con} \\ &+ a\int \mathrm{d}^4y\;\langle\phi_0(x_1)\ldots\phi_0(x_n)\mathcal{L}_1(y)\rangle_{\rm con} \\ &+ a\sum_{k=1}^{n}\langle\phi_0(x_1)\ldots\phi_1(x_k)\ldots\phi_0(x_n)\rangle_{\rm con} + \mathrm{O}(a^2) \end{aligned}$$

- $\langle\cdots\rangle$ is defined w.r.t. continuum theory with S_0
- the a-dependence is now explicit, up to logarithms, which are hidden in the coefficients $c_{1,i}$ and
- In perturbation theory one expects to l-loop order the asymptotic expansion:

$$P(a) \sim P(0) + \sum_{n=1}^{\infty}\sum_{k=0}^{l} g_0^{2k}\, p_{nk}\, a^n (\ln a)^k$$

where e.g. $P(a) = G_n$ at fixed arguments.

Symanzik's effective continuum theory (4)

Conclusions from Symanzik's analysis:

- Asymptotically, cutoff effects are powers in a, modified by logarithms;
- In contrast to Wilson quarks, only even powers of a are expected for
  - 4-dim. bosonic theories (pure gauge theories, scalar field theories)
  - 4-dim. fermionic theories which retain a remnant axial symmetry (overlap, Domain Wall Quarks, staggered quarks, Wilson quarks with a twisted mass term,...)

In QCD simulations a is typically varied by a factor 2

$\Rightarrow$ logarithms vary too slowly to be resolved; linear or quadratic fits (in a resp. a^2) are used in practice.

Symanzik improvement (1)

Improved action and fields:

- Idea: introduce lattice representatives of the operators in $\mathcal{L}_{\mathrm{k}}$ and ϕ_k and include them in the lattice action and fields.

⇒ if coefficients are chosen appropriately the O(a effects can be made to vanish, the lattice action and fields are O(a) improved ($c_{1i} = 0$ in effective action),

- Main problem: how to choose the improvement coefficients?
  - compute the coefficients in lattice perturbation theory; always possible (if tedious), but improvement is not complete.
  - try to determine the coefficients non-perturbatively (s. below)
- Simplification by restriction to on-shell quantities:
  - spectral quantities (particle masses, energies)
  - correlation functions $G_n(x_1, x_2, \ldots, x_n)$, with $x_i \neq x_j$.

  ⇒ use equations of motion to reduce the number of O(a) counterterms.
  - Wilson quarks to O(a); eliminate 2 counterterms, stay with

$$\overline{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi, \qquad m^2\overline{\psi}\psi, \qquad m\,\mathrm{tr}\{F_{\mu\nu}F_{\mu\nu}\}.$$

Symanzik improvement (2)

1. On-shell O(a) improved lattice QCD action
   - The last two terms are equivalent to a rescaling of the bare mass and coupling ($m_\mathrm{q} = m_0 - m_\mathrm{cr}$):

$$\tilde{g}_0^2 = g_0^2(1 + b_g(g_0)am_\mathrm{q}), \qquad \tilde{m}_\mathrm{q} = m_\mathrm{q}(1 + b_\mathrm{m}(g_0)am_\mathrm{q})$$

   - The only new structure is the Sheikholeslami-Wohlert or clover term

$$S_\mathrm{Wilson} \rightarrow S_\mathrm{Wilson} + \tfrac{i}{4} a c_\mathrm{sw}(g_0^2) a^4 \sum_x \overline{\psi}(x)\sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi(x)$$

2. On-shell O(a) improved axial current and density:

$$\begin{aligned}
(A_\mathrm{R})_\mu^a &= Z_\mathrm{A}(\tilde{g}_0{}^2)(1 + b_\mathrm{A}(g_0)am_\mathrm{q})\left\{A_\mu^a + c_\mathrm{A}(g_0)\tilde{\partial}_\mu P^a\right\} \\
(P_\mathrm{R})^a &= Z_\mathrm{P}(\tilde{g}_0{}^2, a\mu)(1 + b_\mathrm{P}(g_0)am_\mathrm{q})P^a
\end{aligned}$$

Determination of improvement coefficients

In principle determine $\mathrm{O}(a)$ improvement coefficients non-perturbatively as functions of g_0^2 by imposing $(i = 1, 2, \ldots)$

$$P_i(a; c_{\mathrm{sw}}, c_{\mathrm{A}}, \ldots) = P_i(0), \quad \Rightarrow c_{\mathrm{sw}}(g_0^2), c_{\mathrm{A}}(g_0^2) \ldots$$

- problem: requires knowledge of the very continuum results $P_i(0)$ which $\mathrm{O}(a)$ improvement should help to obtain!
- can be done in perturbation theory
- Observation: $\mathrm{O}(a)$ counterterms with Wilson fermions arise due to explicit breaking of chiral symmetry!

$\Rightarrow$ can be determined by imposing chiral symmetry relations at finite lattice spacing (i.e. fixed g_0).

Improvement conditions from chiral symmetry

- In QFT symmetries are expressed by Ward identities
- introduce axial field variations:

$$\delta_{\rm A}^{a}(\theta)\psi(x) = i\gamma_5 \tfrac{1}{2}\tau^a\theta(x)\psi(x), \qquad \delta_{\rm A}^{a}(\theta)\overline{\psi}(x) = \overline{\psi}(x) i\gamma_5 \tfrac{1}{2}\tau^a\theta(x)$$

with $\theta(x) = 1$ if $x \in R$ and $= 0$ otherwise ($R =$ space-time region)

- Derive continuum Ward identities by assuming that the functional integral can be treated like an ordinary integral:

$$\begin{aligned}
\langle \delta_{\rm A}^{a}(\theta) O \rangle &= \langle O \delta_{\rm A}^{a}(\theta) S \rangle, \\
\delta_{\rm A}^{a}(\theta) S &= -i \int {\rm d}^4 x \theta(x) \left(\partial_\mu A_\mu^a(x) - 2mP^a(x)\right) \\
A_\mu^a(x) &= \overline{\psi}(x)\gamma_\mu\gamma_5 \tfrac{1}{2}\tau^a\psi(x), \qquad P^a(x) = \overline{\psi}(x)\gamma_5\tfrac{1}{2}\tau^a\psi(x)
\end{aligned}$$

Simplest chiral Ward identity: the PCAC relation

- Choose fields $O = O_{\rm ext}$ located outside region R, then shrink R to a point x:

$$0 = \langle O_{\rm ext}\delta_{\rm A}^a(\theta)S\rangle = -i\left\langle O_{\rm ext}\left(\partial_\mu A_\mu^a(x) - 2mP^a(x)\right)\right\rangle,$$

- Rewritten in terms of the PCAC mass:

$$m = \frac{\langle O_{\rm ext}\partial_\mu A_\mu^a(x)\rangle}{2\langle O_{\rm ext}P^a(x)\rangle}$$

- NOTE: chiral symmetry implies that the PCAC mass must be independent of
  - the choice of $O_{\rm ext}$,
  - the point x or
  - any other parameters such as the space-time volume or shape!

Non-perturbative determination of $c_{\rm sw}$

- There are 2 improvement coefficients in the massless theory ($c_{\rm sw}, c_{\rm A}$), the remaining ones ($b_g, b_{\rm m}, b_{\rm A}, b_{\rm P}$) come with $am_{\rm q}$.
- All counterterms are absent in chirally symmetric regularisations

$\Rightarrow$ turn this around: impose chiral symmetry to determine $c_{\rm sw}, c_{\rm A}$ non-perturbatively:

  - define bare PCAC quark masses from SF correlation functions (i.e. choice of $O_{\rm ext}$ etc.):

$$m_{\rm R} = \frac{Z_{\rm A}(1 + b_{\rm A}\, am_{\rm q})}{Z_{\rm P}(1 + b_{\rm P}\, am_{\rm q})} m, \qquad m = \frac{\tilde\partial_0 f_{\rm A}(x_0) + c_{\rm A}\, a\partial_0^*\partial_0 f_{\rm P}(x_0)}{f_{\rm P}(x_0)}$$

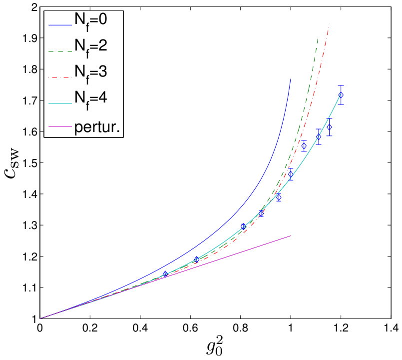
  - At fixed g_0 and $am_{\rm q} \approx 0$ define 3 bare PCAC masses $m_{1,2,3}$ (e.g. by varying the gauge boundary conditions) and impose

$$m_1(c_{\rm sw}, c_{\rm A}) = m_2(c_{\rm sw}, c_{\rm A}), \qquad m_1(c_{\rm sw}, c_{\rm A}) = m_3(c_{\rm sw}, c_{\rm A}) \Rightarrow c_{\rm sw}, c_{\rm A}$$

SF b.c.'s $\Rightarrow$ high sensitivity to $c_{\rm sw}$ & simulations near chiral limit

Results for $c_{\mathrm{sw}}(g_0^2)$

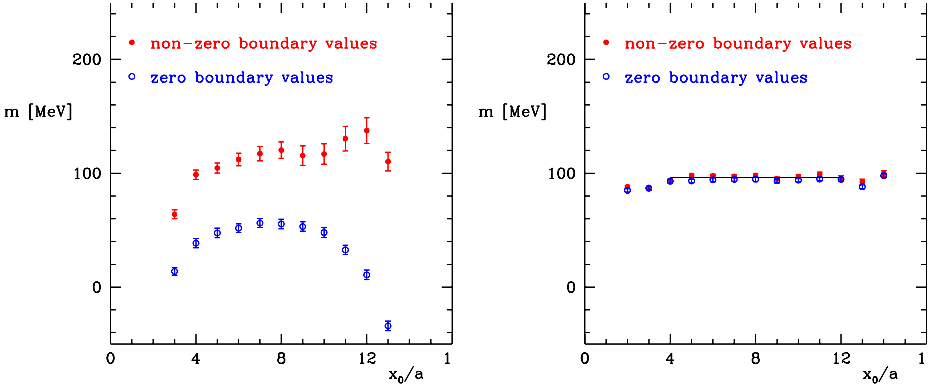
Numerical results in QCD have been obtained for $N_f = 0, 2, 3, 4$ & various gauge actions.



[ALPHA, coll., Sommer, Tekin, Wolff '09]

Effect of on-shell O(a) improvement

Before and after O(a) improvement (PCAC masses from SF correlation functions, $8^3 \times 16$ lattice)



Quenched result for the charm quark mass [ALPHA '02]

- The RGI charm quark mass can be defined in various ways
  - starting from the subtracted bare quark mass $m_{\rm q,c} = m_{0,c} - m_{\rm cr}$
  - starting from the average strange-charm PCAC mass m_{sc}
  - starting from the PCAC mass m_{cc} for a hypothetical mass degenerate doublet of quarks.
- Tune the bare charm quark masses to match the D_s meson mass
- Obtain the corresponding O(a) improved RGI masses:

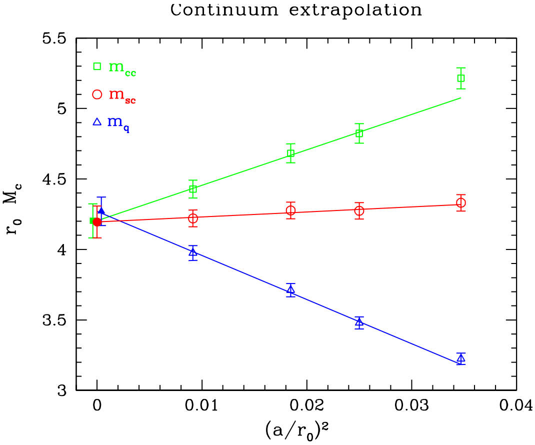
$$\begin{aligned} r_0 M_c|_{m_{sc}} &= Z_M \Big\{ 2 r_0 m_{sc} \left[1 + (b_{\rm A} - b_{\rm P})\tfrac{1}{2}(am_{\rm q,c} + am_{\rm q,s})\right] \\ &\quad - r_0 m_{\rm s} \left[1 + (b_{\rm A} - b_{\rm P}) am_{\rm q,s})\right] \Big\}, \\ r_0 M_c|_{m_{\rm c}} &= Z_M r_0 m_{\rm c} \left[1 + (b_{\rm A} - b_{\rm P}) am_{\rm q,c}\right], \\ r_0 M_c|_{m_{\rm q,c}} &= Z_M Z r_0 m_{\rm q,c} \left[1 + b_{\rm m} am_{\rm q,c}\right]. \end{aligned}$$

- N.B.: all O(a) counterterms are known non-perturbatively in the quenched case!

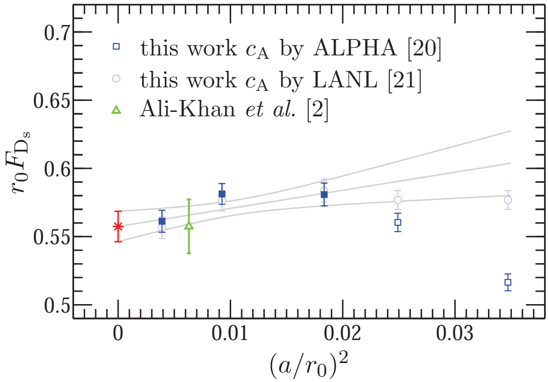
Continuum extrapolation of the quenched RGI charm quark mass

Continuum extrapolation:

$$
\begin{aligned}
r_0 M_c &= A + B(a^2/r_0^2) \\
r_0 &= 0.5\,\mathrm{fm} \\
M_c &= 1.654(45)\,\mathrm{GeV} \\
\overline{m}_c^{\overline{\mathrm{MS}}}(\overline{m}_c) &= 1.301(34)\,\mathrm{GeV}
\end{aligned}
$$



[Heitger, Jüttner '08]



Warning: if a is not small enough the continuum extrapolations can be misguided!

The 2d O(N) sigma model: a test laboratory for QCD?

$$S = \frac{1}{g_0^2} \sum_{x,\mu} (\partial_\mu \mathbf{S})^2, \qquad \mathbf{S} = (S_1, \ldots, S_N) \qquad \mathbf{S}^2 = 1$$

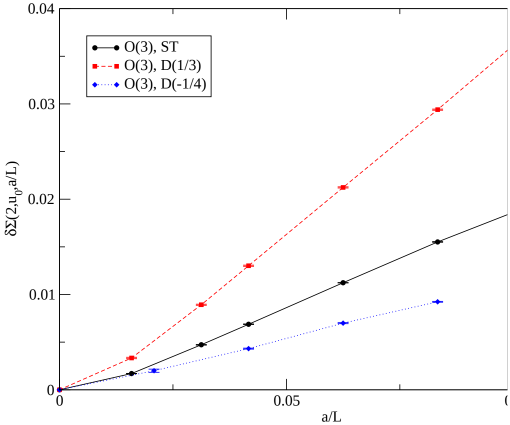
- like QCD the model has a mass gap and is asymptotically free
- many analytical tools: large N expansion, Bethe ansatz, form factor bootstrap;
- some exact results available in continuum limit!
- efficient numerical simulations due to cluster algorithms.

$\Rightarrow$ very precise data over a wide range of lattice spacing (a can be varied by 1-2 orders of magnitude).

- Symanzik: expect O(a^2) effects, up to logarithms!

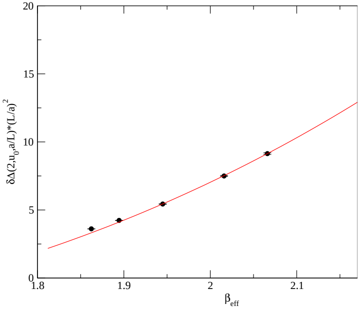
A sobering result:

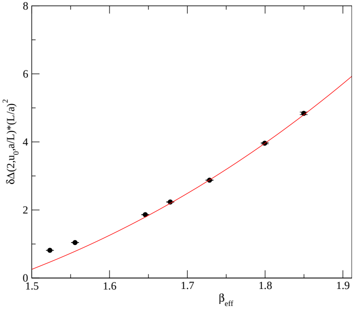
Numerical study of renormalised finite volume coupling to high precision ($N = 3$) [Hasenfratz, Niedermayer '00, Hasenbusch et al. '01, Balog, Niedermayer, Weisz '09]



A closer look:

[Balog, Niedermayer, Weisz '09] cutoff effects $\times L^2/a^2$:





Cuprit: terms like $a^2 \ln^3(a^2)$ conspire to fake a linear behaviour in a over wide range! However, can be understood within Symanzik's effective theory!

O(a^2) effects in lattice QCD

- $\mathcal{L}_2$ contains a host of dimension 6 operators (4-quark operators!) [Sheikholeslami-Wohlert '86]

⇒ complete elimination of O(a^2) effects in lattice QCD à la Symanzik is unpractical!

- Pure gauge theories: O(a^2) terms can be eliminated by adding more extended Wilson loops to the Wilson plaquette action [Lüscher, Weisz '84]; in QCD this is sometimes done for other reasons than O(a^2) improvement.
- O(a) improvement for Wilson quarks gets complicated
  - if quarks are not mass-degenerate
  - for 4-quark operators due to proliferation of O(a) counterterms.

⇒ introduce "twisted mass terms" [Frezzotti, Grassi, S. Weisz, '01] and rely on "automatic O(a) improvement" [Frezzotti, Rossi '03].

Conclusions

- Symanzik's analysis seems applicable beyond perturbation theory; cutoff effects are organized powers of a up to (slowly varying) logarithms
- In lattice QCD numerical results seem to confirm expectations;
- However, large powers of logarithms are a possibility (cf. O(3) model) and could be problematic!
- While Symanzik's effective theory allows to unveal this behaviour, the calculations are tedious! [Balog, Niedermayer, Weisz '09]
- Improvement coefficients are difficult determine non-perturbatively, except where continuum symmetries are broken (e.g chiral symmetry with Wilson quarks)
- Perturbative determinations: automated perturbative methods or stochastic perturbation theory might enable two-loop results which may be sufficient in many cases.