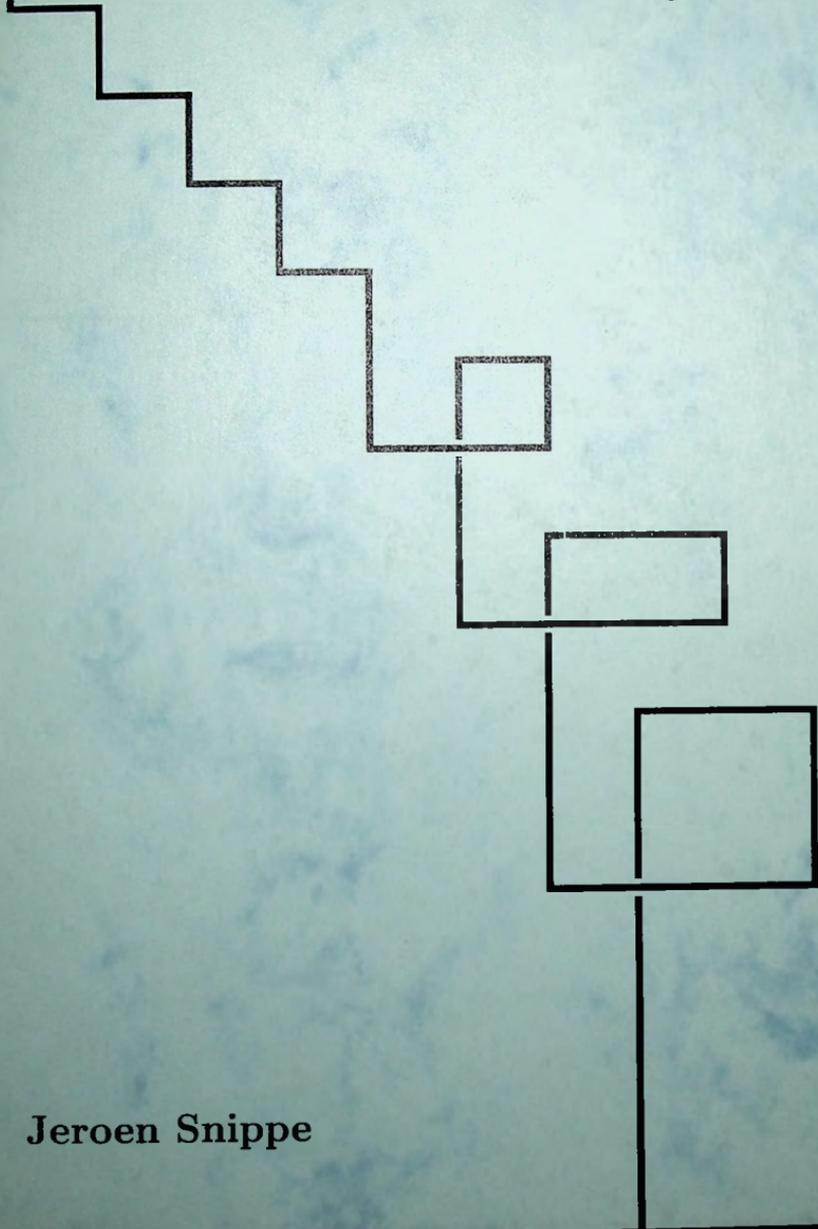


The Uses of  
Improved Actions in  
Lattice Gauge Theory



Jeroen Snippe

1 FEB. 1997

RIJKSUNIVERSITEIT TE LEIDEN  
BIBLIOTHEEK INSTITUUT-LORENTZ  
Postbus 9506 - 2300 RA Leiden  
Nederland

Kast dissertaties

The Uses of  
Improved Actions in  
Lattice Gauge Theory

The loss of  
the world's  
treasure

of the  
ancient world

# The Uses of Improved Actions in Lattice Gauge Theory

Proefschrift

ter verkrijging van de graad van Doctor  
aan de Rijksuniversiteit te Leiden,  
op gezag van de Rector Magnificus Dr. W.A. Wagenaar,  
hoogleraar in de faculteit der Sociale Wetenschappen,  
volgens besluit van het College van Dekanen  
te verdedigen op woensdag 19 februari 1997  
te klokke 15.15 uur

door

Jeroen Robert Snippe

geboren te Leusden op 23 februari 1969

## Promotiecommissie

- Promotor: prof.dr. P.J. van Baal  
Referent: prof.dr. J. Smit (Universiteit van Amsterdam)  
Overige leden: prof.dr. A. González-Arroyo (Universidad Autónoma de Madrid)  
prof.dr. L.J. de Jongh  
prof.dr. R.H.P. Kleiss (Katholieke Universiteit Nijmegen)  
dr. W.L.G.A.M. van Neerven  
prof.dr. B.Q.P.J. de Wit (Universiteit Utrecht)

Voor het gebruik van de Cray supercomputers bij het onderzoek beschreven in de hoofdstukken 2 en 4 van dit proefschrift, is subsidie verleend door de Stichting Nationale Computer Faciliteiten (NCF), met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).



# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Perspective	9
1.2	Lattice gauge theory	14
1.3	Periodic boundary conditions and twist	21
1.4	Outline	26
<b>2</b>	<b>Instantons from over-improved cooling</b>	<b>29</b>
2.1	Introduction	29
2.2	On the existence of continuum solutions	30
2.3	The lattice actions and cooling	32
2.4	Lattice artefacts	33
2.5	Non-leading lattice artefact corrections	37
2.6	Numerical results and discussion	40
<b>3</b>	<b>The <math>O(3)</math> <math>\sigma</math>-model on a cylinder</b>	<b>45</b>
3.1	Introduction	45
3.2	The $O(3)$ $\sigma$ -model in general coordinates	46
3.3	Sphalerons, vacua and other static solutions	48
3.4	Instanton solutions	49
3.4.1	Construction of the solutions	49
3.4.2	Physical interpretation of the moduli space	51
3.5	Conclusion	53
<b>4</b>	<b>Improvement in finite volumes</b>	<b>55</b>
4.1	Introduction	55
4.2	Background fields on the lattice	56
4.3	Abelian potential for the square action	58
4.4	Monte Carlo results	62
4.5	Discussion	63
<b>5</b>	<b>One-loop Symanzik coefficients</b>	<b>65</b>
5.1	Introduction	65
5.2	On-shell improvement	67
5.3	Lattice perturbation theory	69
5.4	Static quark potential	70
5.4.1	Generalities	70
5.4.2	Particulars	75
5.5	Spectroscopy in a twisted finite volume	81

---

5.5.1	Introduction and formalism . . . . .	81
5.5.2	Mass of the $A^+$ meson . . . . .	83
5.5.3	Effective coupling constant . . . . .	89
5.6	Summary . . . . .	95
	Appendix A: Structure of the Feynman rules . . . . .	96
	Appendix B: Propagator . . . . .	100
	Appendix C: Vertex components . . . . .	102
	Appendix D: Feynman rules with a twist . . . . .	107
	Appendix E: Positivity of the square action . . . . .	109
<b>6</b>	<b>Conclusion</b> . . . . .	<b>111</b>
	<b>References</b> . . . . .	<b>113</b>
	<b>List of publications</b> . . . . .	<b>117</b>
	<b>Samenvatting</b> . . . . .	<b>119</b>
	<b>Curriculum Vitae</b> . . . . .	<b>125</b>

1	Introduction	1
2	Chapter I	1
3	Chapter II	1
4	Chapter III	1
5	Chapter IV	1
6	Chapter V	1
7	Chapter VI	1
8	Chapter VII	1
9	Chapter VIII	1
10	Chapter IX	1
11	Chapter X	1
12	Chapter XI	1
13	Chapter XII	1
14	Chapter XIII	1
15	Chapter XIV	1
16	Chapter XV	1
17	Chapter XVI	1
18	Chapter XVII	1
19	Chapter XVIII	1
20	Chapter XIX	1
21	Chapter XX	1
22	Chapter XXI	1
23	Chapter XXII	1
24	Chapter XXIII	1
25	Chapter XXIV	1
26	Chapter XXV	1
27	Chapter XXVI	1
28	Chapter XXVII	1
29	Chapter XXVIII	1
30	Chapter XXIX	1
31	Chapter XXX	1
32	Chapter XXXI	1
33	Chapter XXXII	1
34	Chapter XXXIII	1
35	Chapter XXXIV	1
36	Chapter XXXV	1
37	Chapter XXXVI	1
38	Chapter XXXVII	1
39	Chapter XXXVIII	1
40	Chapter XXXIX	1
41	Chapter XL	1
42	Chapter XLI	1
43	Chapter XLII	1
44	Chapter XLIII	1
45	Chapter XLIV	1
46	Chapter XLV	1
47	Chapter XLVI	1
48	Chapter XLVII	1
49	Chapter XLVIII	1
50	Chapter XLIX	1
51	Chapter L	1
52	Chapter LI	1
53	Chapter LII	1
54	Chapter LIII	1
55	Chapter LIV	1
56	Chapter LV	1
57	Chapter LVI	1
58	Chapter LVII	1
59	Chapter LVIII	1
60	Chapter LIX	1
61	Chapter LX	1
62	Chapter LXI	1
63	Chapter LXII	1
64	Chapter LXIII	1
65	Chapter LXIV	1
66	Chapter LXV	1
67	Chapter LXVI	1
68	Chapter LXVII	1
69	Chapter LXVIII	1
70	Chapter LXIX	1
71	Chapter LXX	1
72	Chapter LXXI	1
73	Chapter LXXII	1
74	Chapter LXXIII	1
75	Chapter LXXIV	1
76	Chapter LXXV	1
77	Chapter LXXVI	1
78	Chapter LXXVII	1
79	Chapter LXXVIII	1
80	Chapter LXXIX	1
81	Chapter LXXX	1
82	Chapter LXXXI	1
83	Chapter LXXXII	1
84	Chapter LXXXIII	1
85	Chapter LXXXIV	1
86	Chapter LXXXV	1
87	Chapter LXXXVI	1
88	Chapter LXXXVII	1
89	Chapter LXXXVIII	1
90	Chapter LXXXIX	1
91	Chapter LXXXX	1
92	Chapter LXXXXI	1
93	Chapter LXXXXII	1
94	Chapter LXXXXIII	1
95	Chapter LXXXXIV	1
96	Chapter LXXXXV	1
97	Chapter LXXXXVI	1
98	Chapter LXXXXVII	1
99	Chapter LXXXXVIII	1
100	Chapter LXXXXIX	1
101	Chapter LXXXXX	1

# 1 Introduction

## 1.1 Perspective

A proton is a theoretically poorly understood object. It is widely accepted that it consists of three quarks, and that the interaction between the quarks is described by Quantum Chromodynamics (QCD). However, the mechanism by which this leads to bound states (hadrons) such as protons and neutrons is a complicated one, in particular at the quantitative level. Much knowledge comes from numerical simulations, but an analytic understanding with a satisfactory predictive power is still incomplete.

QCD is one of the two ingredients of the successful Standard Model of high energy physics. It describes the strong force, i.e. the force between protons and neutrons, or, more fundamentally, between quarks. Like in the electroweak sector of the Standard Model, the strong force itself is mediated by particles, called gluons. It is only at high energies that QCD is well understood, as it has proven to give accurate predictions for scattering phenomena in particle accelerators.

QCD is a gauge theory similar to Quantum Electrodynamics (QED), with quarks corresponding to electrons, and gluons to the quanta of the electromagnetic field, photons. However, there is an important difference between the two theories. Gluons have direct mutual interactions, while photons do not. At small distances (or by a Fourier transformation, at high momenta) the interactions between gluons are small. This is what is called asymptotic freedom [1], and it is for this reason that we can calculate phenomena in particle accelerators by means of a perturbative expansion. However, at larger distances the coupling grows, and at some point (typically the size of light hadrons) it is so large that perturbation theory breaks down. It is for large distances that QCD behaves vastly different than QED.

In particular, QCD exhibits quark confinement. This is the phenomenon that quarks cannot exist as free particles, but only appear in colorless bound states (color is the analogue of electric charge in QED). Originally, when QCD was invented to account for the structure of hadrons, it was a postulate that confinement is realized in QCD, necessary because free quarks have never been observed. Over the past years, Monte Carlo simulations have brought convincing numerical evidence for the validity of this postulate. The simulations support the picture that a string of gluons forms between (anti) quarks when their distance is large enough. This string exerts a constant force on the quarks (called the string tension). Equivalently, the energy carried by the gluons in the string grows linearly with the distance. It is only at very high temperature (roughly  $10^{12}$  K) that QCD undergoes a deconfining phase transition, at which a quark-gluon plasma is formed. In this thesis we limit ourselves to zero temperature.

Because gluons play such an important role in the large-distance behavior of

QCD, it is sensible, at least as a first step, to leave out the quarks altogether. The resulting model is called 'pure gauge theory'. An extra justification for doing pure gauge theory, is the existence of glueballs. These are hadrons supposedly consisting mainly out of gluons. Recent simulations [2] indicate a rather small width of 0.1 GeV for the pure gauge scalar glueball coupled to the meson decay channels, seemingly justifying the large gluon content. The authors present arguments for identification with an observed resonance at 1.7 GeV. Other groups [3] claim the scalar glueball rather at 1.5 GeV.

Pure gauge theory is the framework of this thesis. We will now discuss it in more detail. The gluons are associated with a non-Abelian gauge group  $G$ , and their dynamics is brought about through the Lagrangian density<sup>1</sup>

$$\begin{aligned} \mathcal{L}(x) &= -\frac{1}{2} \sum_{\mu, \nu} \text{Tr} F_{\mu\nu}^2(x), \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu], \quad A_\mu = g_0 \sum_a A_\mu^a T^a, \end{aligned} \quad (1.1.1)$$

where  $g_0$  is the (bare) coupling constant,  $T^a$  are the generators of  $G$  (in the fundamental representation) and  $A_\mu^a(x) \in \mathbb{R}$ . In principle one should take  $G = \text{SU}(3)$  as in QCD, but often the choice  $G = \text{SU}(2)$  is made. This is expected not to change the qualitative behavior of the model, while it simplifies calculations. In this thesis, the choice of  $G$  will always be restricted to  $\text{SU}(N)$ , except for small excursions to  $\text{U}(1)$ . An essential property of the Lagrangian (1.1.1) is that it is invariant under local gauge transformations

$$A_\mu(x) \rightarrow A_\mu^\Omega(x) \equiv \Omega(x)(A_\mu(x) + \partial_\mu)\Omega^{-1}(x), \quad \Omega(x) \in G. \quad (1.1.2)$$

For gauge theories it is difficult to do analytic non-perturbative calculations in a reliable way. The reason is that gauge theory, being a field theory, has an infinite number of degrees of freedom, one for each field component at each point in space-time. All these degrees of freedom interact with one another, and the only known way to keep track of the interactions is perturbation theory. However, *numerically* non-perturbative calculations can be done by means of Monte Carlo simulations. Also here the infinite number of degrees of freedom, simply not fitting in a computer, poses a problem, but it can be overcome by making space-time discrete and finite. Usually one chooses a hypercubic lattice with lattice spacing  $a$ , in a space-time  $[0, L]^3 \times [0, T]$  with periodic boundary conditions.

This choice breaks Lorentz invariance<sup>2</sup>, but the fundamental property of gauge invariance can be maintained by choosing an appropriate lattice action (we will come back to this in section 1.2). Depending on  $g_0$ , the resulting discretized model has a certain correlation length  $\xi$ . The continuum limit is obtained by taking  $a/\xi \rightarrow 0$ , which due to asymptotic freedom corresponds to  $g_0 \rightarrow 0$ , while Lorentz invariance should be restored for  $L/\xi \rightarrow \infty, T/\xi \rightarrow \infty$ . The existence of a continuum limit is

<sup>1</sup>We will usually formulate the field theory in Euclidean space-time.

<sup>2</sup>More accurately, rotational invariance in Euclidean four-dimensional space-time.

expected from (perturbative) renormalizability of pure gauge theory [4], and should go through for a wide class of (gauge invariant) lattice actions due to universality arguments [5]. Monte Carlo results support these expectations. To all orders in perturbation theory, renormalizability and universality of lattice gauge theories was proven in ref. [6].

In Monte Carlo simulations,  $L/a$  and  $T/a$  are finite integer numbers, and therefore systematic errors are made. Finite volume errors behave as  $\exp(-L/\xi)$ , and thus quickly drop to zero. In this thesis we will consider lattice errors (or lattice artefacts), i.e. errors due to the finiteness of  $\xi/a$ . For the standard lattice action, called the Wilson action [7], these are of the order  $(a/\xi)^2$ . However, improved lattice actions exist that are constructed in such a way as to make the artefacts systematically smaller. This is discussed in more detail in the next section.

We now turn to the feasibility of *analytic* non-perturbative calculations, in particular with respect to the mass spectrum. In small spatial volumes reliable calculations are possible. The reason is that, again due to asymptotic freedom, the renormalized coupling constant  $g_R(L)$  vanishes for  $L \rightarrow 0$ , so that physical quantities have a well-behaved perturbative expansion with respect to  $g_R(L)$ . The interesting point is that by increasing  $L$ , the onset of non-perturbative phenomena can be studied, and even included in the computation [8]. To make this more precise it is best to use the Hamiltonian formalism, which can be obtained from the Lagrange formulation by choosing the gauge  $A_0 = 0$  and performing a Legendre transformation in the usual way. The result reads

$$\begin{aligned} \hat{H} &= \frac{1}{2} \int d^3\mathbf{x} \sum_{i,a} \hat{\Pi}_i^a(\mathbf{x})^2 + \hat{V}, \quad \hat{V} = \frac{1}{2} \int d^3\mathbf{x} \sum_{i,a} \hat{B}_i^a(\mathbf{x})^2, \\ \hat{B}_i^a &= \frac{1}{2} \sum_{j,k} \varepsilon_{ijk} \hat{F}_{jk}^a, \quad \hat{F}_{ij}^a = \partial_i \hat{A}_j^a - \partial_j \hat{A}_i^a + g_0 \sum_{b,c} f_{abc} \hat{A}_i^b \hat{A}_j^c, \end{aligned} \quad (1.1.3)$$

with  $[\hat{\Pi}_i^a(\mathbf{x}), \hat{B}_k^b(\mathbf{y})] = -i\delta_{ij}\delta_{ab}\delta^3(\mathbf{x} - \mathbf{y})$ . As configuration space associated with this Hamiltonian one should take  $\mathcal{A}/\mathcal{G}$ , where  $\mathcal{A}$  is the set of all allowed fields  $A_i^a(\mathbf{x})$ , and  $\mathcal{G}$  is the set of spatial gauge transformations<sup>3</sup>.

The analysis of the spectrum of  $\hat{H}$  is complicated by the fact that  $\mathcal{A}/\mathcal{G}$  is infinite dimensional, and furthermore topologically highly non-trivial. In small volumes, however, both complications are manageable. The reason is that in that case wave functions on  $\mathcal{A}$  are concentrated in the classical vacua, i.e. points in  $\mathcal{A}$  where the potential  $V$  takes its minimal value 0. It is only for larger values of  $L$ , i.e. larger values of  $g_R(L)$ , that the wave functions start spreading out over the barrier surrounding the classical vacua. This happens first for states with high energies, but for large enough values of  $L$  also low-lying states will be affected.

Let us concentrate on the situation where only periodic fields  $A_i(\mathbf{x})$  are included<sup>4</sup>,

<sup>3</sup>In this formulation the kinetic term in eq. (1.1.3) actually receives corrections due to the non-trivial metric on  $\mathcal{A}/\mathcal{G}$  [9], but these are not essential to the discussion below.

<sup>4</sup>The analysis is different if one allows for twist [10], where only gauge invariant fields are required to be periodic. We will come back to this in section 1.3.

for  $G = \text{SU}(2)$ . In that case the set of (classical) vacua, also called the vacuum valley, is 3-dimensional. In terms of  $A_i(\mathbf{x})$  it can be parametrized by

$$A_j(\mathbf{x}) = \frac{1}{2} i \sigma_3 \frac{C_j}{L}, \quad (1.1.4)$$

where  $C_j \in \mathbb{R}$ . Due to the possibility of periodic gauge transformations,  $C_j = 4\pi k_j$  are to be identified to one another ( $k_j \in \mathbb{Z}$ ), and also  $\pm(C_1, C_2, C_3)$  represent the same point. For a derivation see ref. [8].

Quantum mechanically, it is possible to take a Born-Oppenheimer approach: integrate out all modes orthogonal to the vacuum valley, keeping the vacuum valley modes (1.1.4) as a background field, and then diagonalize the resulting effective Hamiltonian for the modes (1.1.4). Actually, in a careful analysis some extra (constant) modes should be kept as a background field<sup>5</sup>. However, the point we want to make here is that the effective Hamiltonian describes ordinary quantum mechanics for a finite number of degrees of freedom. It can therefore be diagonalized by conventional methods. For the lowest-lying states this approach works well [8, 11] for  $\xi L \lesssim 5$ , as comparison with accurate Monte Carlo data [12] has shown convincingly.

For larger values of  $L$ , the method breaks down due to the spreading of the wave functions. However, in principle this effect can be taken into account too, once one realizes that the spreading over the barrier surrounding the vacuum valley takes place predominantly at the point (or points) where the barrier is lowest (such points are called sphalerons [13]). By *not* integrating out sphaleron directions, but including them in the effective Hamiltonian, the spreading can thus be incorporated non-perturbatively (at least up to somewhat larger values of  $L$ ). The topological nature of  $A/G$  comes into play here, because the loop 'vacuum valley  $\rightarrow$  sphaleron  $\rightarrow$  vacuum valley' (where the latter vacuum valley is on the other side of the barrier) is non-contractable. In fact, the mathematical definition of a sphaleron [14], by a minimax procedure, is based on this topological structure: For a non-contractable loop starting and ending in the vacuum valley one determines the maximal value of the potential  $V$  for configurations along the loop. Then one minimizes this value over all such loops. The resulting value is called the sphaleron energy, and the corresponding configuration(s) is (are) called sphaleron(s). For a simple example see fig. 1-1. One can prove that any sphaleron is a saddle point of  $V$ , with only one unstable mode. If the set of sphalerons, or sphaleron moduli-space, is continuous, there are of course also zero modes.

The extension discussed in the previous paragraph was recently completed [15] (see also ref. [16]) for  $\text{SU}(2)$  gauge theory on a spatial three-sphere  $S^3$ , instead on the three-torus  $[0, L]^3$ .  $S^3$  is better suited for doing analytic calculations than  $[0, L]^3$ , due to its larger symmetry group. Nevertheless, it would be important to repeat the calculation for  $[0, L]^3$ , because for this geometry results can be compared to Monte Carlo data. Unfortunately for  $[0, L]^3$  the sphaleron configurations are not known analytically.

<sup>5</sup>The reason is that for  $C_j = 0 \pmod{2\pi}$  the potential  $V$  rises quartically in the direction of the extra constant modes, rather than quadratically.

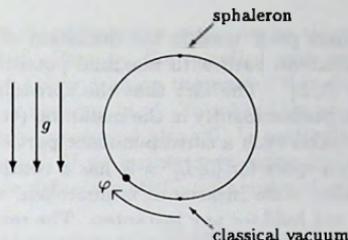


Figure 1-1. A point particle on the unit circle in a homogeneous force field  $g$ . Its potential is  $V(\varphi) = g(1 - \cos \varphi)$ . The classical vacuum is  $\varphi = 0$ , and the sphaleron is  $\varphi = \pi$ , while an instanton makes one full turn from  $\varphi = 0$  to  $\varphi = 2\pi$  (see the text for definitions). Quantum states with energies much smaller than  $2g$  can be accurately described in perturbation theory around  $\varphi = 0$ . At somewhat higher energies, tunneling transitions  $\varphi = 0 \rightarrow 2\pi$  must be taken into account, and for still higher energies wave functions will spread over the whole circle.

We end this section with a discussion on the relation between sphalerons and instantons [17]. Such a relation may be used to simplify the search for sphalerons, as in a direct approach instantons often are easier to find than sphalerons.

An instanton is a (classical) field configuration  $A_\mu(x)$  in Euclidean space-time, that can only exist if the system {gauge group, space-time} is of topologically non-trivial nature. In that case the field space  $\{A_\mu(x)\}$  can be divided in topological sectors characterized by winding numbers  $\nu \in \mathbb{Z}$ . The minimal Euclidean action in a sector is  $8\pi|\nu|$ . An instanton is a minimal-action configuration in the sector  $\nu = 1$  (for  $\nu = 0$  the minimum is  $A_\mu(x) = 0$ , or a gauge transformation thereof). An important property of the minima in any sector is that they not only satisfy the (Euclidean) Euler-Lagrange equations, but also the more restrictive self-duality equations,  $E_i^a = B_i^a$  (where  $E_i \equiv F_{0i}$ ). Usually the set of instantons, also referred to as the instanton moduli-space, has non-zero dimension. In fact, for compact spacetimes an index theorem is applicable which states that the moduli-space in the  $\nu^{\text{th}}$  sector is  $8\nu - n$  dimensional, where  $n$  depends on the geometry (for example, for  $S^4$ ,  $n = 3$ , while for  $T^4$ ,  $n = 0$ ). These moduli can often be understood in terms of symmetries of the action. The physical importance of instantons lies in their relevance to semiclassical expansions of the path integral [18, 19]. Analytic expressions for instantons have been found for the space-time  $S^4$ , from which instantons on  $\mathbb{R}^4$  and  $S^3 \times \mathbb{R}$  can be obtained by conformal transformations.

Let us assume that instantons also exist<sup>6</sup> on  $[0, L]^3 \times \mathbb{R}$ . Since, in the gauge  $A_0 = 0$ , the Euclidean action reads  $S = -\frac{1}{2} \int d^4x \text{Tr} E_i^a{}^2(x) + \int dt V(t)$ , it is clear from the finiteness of the instanton action that an instanton must have endpoints  $A_i(\mathbf{x}, t \rightarrow \pm\infty)$  in the vacuum valley. Furthermore, an instanton, viewed as a path in  $A/G$  parametrized by  $t$ , is non-contractable. Therefore an instanton path is a good

<sup>6</sup>In fact, the search for such instantons will be the subject of chapter 2.

candidate for the 'mini-max path' used in the definition of the sphaleron, in which case the point on the instanton path with maximal potential would be a sphaleron on the spatial geometry  $[0, L]^3$ . The fact that the spreading of the wave function, mentioned earlier, will be predominantly in the instanton (tunneling) directions when the barrier is still high, makes such a correspondence particularly natural.

If the instanton moduli-space for  $[0, L]^3 \times \mathbb{R}$  has a continuous scale parameter  $\rho$ , as is the case for  $\mathbb{R}^4$  where scale invariance is unbroken, the forementioned correspondence will certainly not hold for any instanton. The reason is that all instantons have equal action, and therefore smaller instantons must cross the barrier at higher energies. From this argument one sees that only the instanton with maximal scale, the existence of which is guaranteed by the finiteness of  $L$ , can possibly go through a sphaleron.

## 1.2 Lattice gauge theory

The main aim in the construction of a lattice action, is the preservation of gauge invariance. This can be achieved through the use of Wilson loops [7], gauge invariant<sup>7</sup> objects associated with closed paths in space-time. The subclass of loops following the links connecting lattice sites, can be considered lattice objects, and can in fact be used for the construction of a suitable lattice action. This construction is the topic of the present section. It will be shown that the lattice action is by no means unique, so that ample room for the reduction of lattice artefacts remains.

To define Wilson loops, it is instructive first to consider the Abelian case  $G = U(1)$ . Extracting a factor  $i$  (the generator of  $U(1)$ ) and writing  $\Omega = \exp(i\Lambda)$ , the gauge transformation (1.1.2) simplifies to  $A_\mu \rightarrow A_\mu - \partial_\mu \Lambda$ . It is easy to make a gauge invariant object out of  $A_\mu$  that is associated with a closed oriented path  $\mathcal{C}$ , namely

$$W_{\mathcal{C}}[A] \equiv \exp \left( i \int_{\mathcal{C}} dx_\mu A_\mu(x) \right). \quad (1.2.1)$$

(Until we come to the lattice formulation, we use the summation convention over repeated indices). Gauge invariance is guaranteed because  $\int_{\mathcal{C}} dx_\mu \partial_\mu \Lambda = 0$ , under the assumption that the loop is contractable<sup>8</sup>. The exponentiation in eq. (1.2.1) is not necessary for gauge invariance, but it is natural because it yields  $W_{\mathcal{C}}[A] \in U(1)$ . Note that by Stokes' theorem,  $\ln W_{\mathcal{C}}[A]$  can also be expressed as a surface integral of  $F_{\mu\nu}$ .

For arbitrary gauge group  $G$ , eq. (1.2.1) generalizes to (with the generators of  $G$  absorbed in  $A_\mu$  as in eq. (1.1.1))

$$W_{\mathcal{C}, \nu, \nu'}[A] \equiv \text{Pexp} \left( \int_{\mathcal{C}} dx_\mu A_\mu(x) \right). \quad (1.2.2)$$

<sup>7</sup>Up to a conjugation, see below.

<sup>8</sup>If space-time is toroidal, non-contractable loops exist, namely loops that wind around the torus. Such loops are only invariant under a subclass of gauge transformations. Operators  $W_{\mathcal{C}}[A]$  for  $\mathcal{C}$  a straight line winding around the torus once, are usually referred to as Polyakov lines.

For later convenience we do not restrict ourselves to closed paths, and denote the endpoints of  $C$  by  $y$  and  $y'$ .  $Pexp \int_C$ , the so-called path ordered exponential, is the product of factors  $\exp(A_\mu(x)dx_\mu)$  along the path  $C$  ( $dx$  being an infinitesimal line segment), starting in  $y$  and ending in  $y'$ . To make this precise, parametrize a point  $x \in C$  by its distance  $s$  to  $y$ , measuring along  $C$ . In particular, if  $D$  is the total length of  $C$ ,  $x(0) = y$  and  $x(D) = y'$ . If we furthermore define  $\hat{x}(s)$  to be the tangent vector of  $C$  at the point  $x(s)$ , the definition of the path ordered exponential reads

$$Pexp\left(\int_C dx_\mu A_\mu(x)\right) \equiv \lim_{M \rightarrow \infty} e^{\varepsilon \hat{x}_\mu(0) A_\mu(x(0))} e^{\varepsilon \hat{x}_\mu(\varepsilon) A_\mu(x(\varepsilon))} e^{\varepsilon \hat{x}_\mu(2\varepsilon) A_\mu(x(2\varepsilon))} \\ \times \dots e^{\varepsilon \hat{x}_\mu(D-\varepsilon) A_\mu(x(D-\varepsilon))}, \quad (\varepsilon \equiv \frac{D}{M}). \quad (1.2.3)$$

It is clear that  $W_{C,y,y'}[A] \in G$ . Also one easily checks that under a gauge transformation (1.1.2),  $\exp[\varepsilon \hat{x}_\mu(s) A_\mu(x(s))] \rightarrow \Omega(x(s)) \exp[\varepsilon \hat{x}_\mu(s) A_\mu(x(s))] \Omega^{-1}(x(s+\varepsilon)) + \mathcal{O}(\varepsilon^2)$ , so that

$$W_{C,y,y'}[A] \rightarrow W_{C,y,y'}^\Omega[A] \equiv W_{C,y,y'}[A^\Omega] = \Omega(y) W_{C,y,y'}[A] \Omega^{-1}(y'). \quad (1.2.4)$$

In particular the Wilson loop  $W_C$  is gauge invariant up to a conjugation, so that its trace is gauge invariant. Also,  $\text{Tr} W_{C,y,y'}$  is independent of the choice of  $y$  on the closed path  $C$ . As a final point note that  $W_{C,y,y'}$  satisfies the differential equations

$$\hat{x}_\mu(0) \mathcal{D}_\mu(y) W_{C,y,y'} = 0 = \hat{x}_\mu(0) \mathcal{D}_\mu(y') W_{C,y,y'}^{-1}, \quad (1.2.5)$$

where  $\mathcal{D}_\mu(x) \equiv \frac{\partial}{\partial x_\mu} + A_\mu(x)$  is the covariant derivative in the fundamental representation.

After these preliminaries we are ready for the lattice formulation<sup>9</sup>. To this end let us superimpose a hypercubic lattice over the continuum. A two-dimensional section through a layer of lattice sites in the  $(\mu, \nu)$  plane is depicted in figure 1-2. In this figure we also included a Wilson loop of a special kind, namely the smallest loop that fits on the lattice. We can use this loop, which will be denoted by  $U_{\mu\nu}(x)$ , to define the Wilson lattice action [7] (the summation convention will be dropped from now on):

$$S_W = \sum_{x,\mu,\nu} \text{Tr} (1 - U_{\mu\nu}(x)). \quad (1.2.6)$$

From the preceding analysis, this action is manifestly gauge invariant. Furthermore, one can expand  $S_W$  with respect to the lattice spacing  $a$ , which is appropriate in view of the continuum limit  $a \rightarrow 0$ . In leading order this leads to

$$S_W = -\frac{1}{2} a^4 \sum_{x,\mu,\nu} \text{Tr} F_{\mu\nu}^2 (1 + \mathcal{O}(a^2)) = -\frac{1}{2} \sum_{\mu,\nu} \int d^4x \text{Tr} F_{\mu\nu}^2 + \mathcal{O}(a^2), \quad (1.2.7)$$

<sup>9</sup>An extensive introduction to lattice gauge theory, including many additional considerations, can be found in ref. [20].

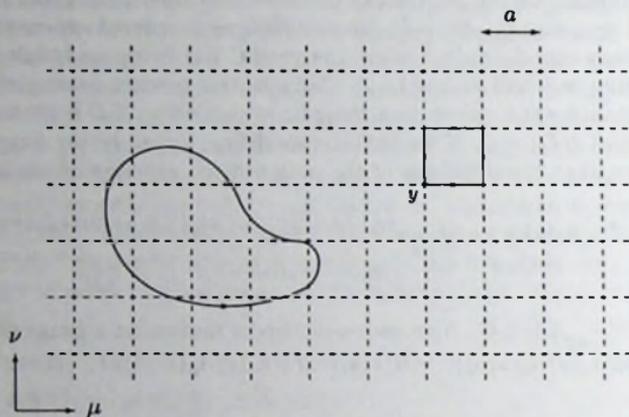


Figure 1-2. A square lattice in the  $(\mu, \nu)$  plane with lattice spacing  $a$ . Two closed paths are depicted. The left one is a generic loop in this plane. The other one is a loop that fits on the lattice, meaning that it consists only of links between neighboring lattice sites. The latter is furthermore special, because it is the smallest possible closed path on the lattice (with non-zero area). We will denote the associated Wilson loop by  $U_{\mu\nu}(y)$ , where  $y$  is the lower left corner point.

the derivation of which can be found in chapter 2. We see that for  $a \rightarrow 0$ ,  $S_W$  reduces to the continuum action defined by eq. (1.1.1), so we can conclude that  $S_W$  is a good discretization of the continuum action.

Of course we have been cheating.  $S_W$  is not yet a true lattice action, because  $U_{\mu\nu}(x)$  is defined in terms of the continuum field  $A_\mu$ . To cure this, we have to define an elementary lattice field. For this purpose we introduce the so-called link variables  $U_\mu(x) \in G$  ( $x$  a lattice site):  $U_\mu(x)$  is associated with the oriented link between the lattice sites  $x$  and  $x + a\hat{\mu}$  (where  $\hat{\mu}$  is the unit vector in the positive  $\mu$ -direction), and will therefore often be denoted by

$$U_\mu(x) = e^{i a A_\mu(x)}. \quad (1.2.8)$$

In principle these elementary lattice variables are completely unrelated to the continuum field, though one is always free to make the *correspondence*

$$U_\mu(x) \leftrightarrow \text{Pexp} \left( \int_0^a ds A_\mu(x + s\hat{\mu}) \right). \quad (1.2.9)$$

If the total number of lattice sites is  $\mathcal{N}$ , then the lattice gauge theory has  $\mathcal{N} \cdot \dim(G)$  degrees of freedom. In terms of these, the small Wilson loop (also called plaquette) is redefined as

$$U_{\mu\nu}(x) \equiv \square_{x, \mu}^\nu \equiv U_\mu(x) U_\nu(x + a\hat{\mu}) U_\mu^\dagger(x + a\hat{\nu}) U_\nu^\dagger(x), \quad (1.2.10)$$

and now the action (1.2.6) has become a genuine lattice action.

As announced, gauge invariance is maintained. The lattice action is invariant under

$$U_\mu(x) \rightarrow U_\mu^\Omega(x) \equiv \Omega(x)U_\mu(x)\Omega^{-1}(x+a\hat{\mu}), \quad \Omega(x) \in G, \quad (1.2.11)$$

where  $\Omega(x) \in G$  is arbitrary at any lattice site  $x$ . The gauge invariance thus is truly local. Note that in view of eqs. (1.2.11) and (1.2.4), the correspondence (1.2.9) is particularly natural.

So now we have a genuine lattice gauge theory. But can we still make contact with the continuum theory? After all, we are not at all forced to read eq. (1.2.9) as an equality.

If one is satisfied with a classical answer, the problem is easily resolved. Classically, the objects of interest are solutions to the Euler-Lagrange equations, i.e. extrema of the action. Consider a solution  $A_\mu(x)$  of the continuum theory. Then use eq. (1.2.9) to transform it into a lattice field. By virtue of eq. (1.2.7), which is applicable if the continuum solution is smooth on the scale of the lattice spacing, it follows that  $S_W$  is extremal at  $U_\mu(x)$ , up to  $\mathcal{O}(a^2)$  corrections. In conclusion, classically eq. (1.2.9) can be used as a map from continuum gauge theory to lattice gauge theory, mapping (smooth) solutions into near-solutions.

In a quantum mechanical context, the answer is less clear-cut. The quantum behavior of lattice gauge theory is defined by expectation values of operators  $\mathcal{O}$ , defined by the path integral<sup>10</sup>

$$\begin{aligned} \langle \mathcal{O} \rangle &\equiv \frac{1}{Z} \int \mathcal{D}U \mathcal{O}(\{U\}) \exp\left(-\frac{1}{g_0^2} S_W[U]\right), \\ Z &\equiv \int \mathcal{D}U \exp\left(-\frac{1}{g_0^2} S_W[U]\right), \quad \mathcal{D}U \equiv \prod_{x,\mu} dU_\mu(x), \end{aligned} \quad (1.2.12)$$

where  $dU_\mu(x)$  is the (gauge invariant) Haar measure on  $G$ . The domain of integration includes configurations that are wildly fluctuating at the scale of one lattice spacing, so the above classical analysis is not at all applicable. In fact, the lattice spacing  $a$  does not even appear in  $Z$ ! (This is no coincidence, but a consequence of the scale invariance of the continuum action).

However,  $a$  is *implicitly* present, because the bare coupling  $g_0$  should be chosen to depend on  $a/\xi$ , where  $\xi$  is the correlation length of the lattice theory. Since  $\xi$  is a physical observable, contact with the continuum is made by tuning  $g_0 \rightarrow 0$ , implying  $a/\xi \rightarrow 0$ . By renormalization group arguments, the expansion (1.2.7) remains sensible in the quantum mechanical context: the  $\mathcal{O}(a^2)$  corrections correspond to so-called irrelevant operators that for  $a/\xi \rightarrow 0$  should not influence the behavior at the physical scale  $\xi$ .

This is what is known as universality. It is generally accepted that one is free to choose any lattice action, as long as it is gauge invariant and its naive (i.e. classical)

<sup>10</sup>Note that for a finite lattice, the path integral is completely well defined, even without the need for gauge fixing (assuming  $G$  to be a compact group, e.g.  $SU(N)$ ).

limit for  $a \rightarrow 0$  is proportional<sup>11</sup> to the continuum action. An infinite number of admissible lattice actions thus exist. One is free to replace the Wilson action (1.2.6) by the sum over an arbitrary number of Wilson loops that fit on the lattice, with arbitrary coefficients.

This freedom can be used to go one step further: the construction of improved lattice actions. These are lattice actions yielding results closer to the continuum limit than results produced by the Wilson action at the same value of  $a$ . For this the coefficients of additional loops are to be chosen judiciously. The remainder of this section is devoted to such actions.

In Monte Carlo simulations, improved actions are to be preferred over the Wilson action, because they need a smaller number of lattice points, and hence less computer time, to obtain results with the same accuracy. We are assuming here that the extra cost of simulating an improved action, unavoidable due to the inclusion of extra loops, does not spoil the gain. This assumption is valid even for actions showing only moderate improvement, because the computational costs in lattice gauge theories grow quickly for decreasing  $a$ . Due to a phenomenon called 'critical slowing down' (see e.g. ref. [21]), this growth is faster than the number of lattice points (which is proportional to  $1/a^4$  in a space-time volume of fixed physical size).

There are two main approaches to improvement. The first is based on Wilson's [5] renormalization group. In this approach the aim is to find a renormalized trajectory (RT) in the space of all possible lattice actions. By definition, this is a trajectory related to the continuum action by transformations that increase the lattice spacing, but leave the physics invariant. A point on an RT thus parametrizes a lattice action that is completely free of lattice artefacts. In spite of some early work by Wilson [22], it seems that for a long time the search for an RT in gauge theories was considered hopeless. However, recently Hasenfratz and Niedermayer succeeded in the semi-analytic construction of an approximate RT for the two-dimensional  $O(3)$   $\sigma$ -model, the generalization of which to four-dimensional gauge theories seems to be viable [23]. It can be shown that their improved action is classically perfect, but to higher order in  $g_0$  there are deviations (though it has been argued that these do not yet appear at one-loop order [23]).

In this thesis, however, we concentrate on the second approach. This approach, due to Symanzik [24], is in principle more straightforward, using perturbation theory in  $g_0$ . To discuss this approach, we introduce some notation. Let  $i$  be an index that parametrizes classes of lattice Wilson loops, e.g.  $i = 0$  for  $1 \times 1$  loops ();  $i = 1$  for  $1 \times 2$  loops (); etc. (also non-planar loops are allowed). Let us only consider lattice actions that respect translational and cubic lattice symmetries,

$$S_{\text{Lat}}(\{c_i(g_0)^2\}) = \sum_i c_i(g_0^2) \sum_{C \in \mathcal{C}_i} \text{Tr} (1 - U(C)), \quad (1.2.13)$$

where  $\sum_{C \in \mathcal{C}_i}$  runs over all lattice loops in the  $i^{\text{th}}$  class (it consists of a sum over lattice sites and sums over Lorentz indices). Since the coefficients are bare quantities, they

<sup>11</sup>The constant of proportionality can be absorbed in the bare coupling  $g_0$ , see eq. (1.2.12).

are allowed to depend on  $g_0$ . Note that in the continuum limit  $a \rightarrow 0$ , the perturbative computation of  $c_i(g_0^2)$  is sensible due to asymptotic freedom.

Symanzik improvement to  $n$ -loop<sup>12</sup> order now amounts to choosing  $c_i(g_0^2) \equiv \sum_{m=0}^{\infty} c_i^{(m)} g_0^{2m}$  in such a way that all physical quantities are free of lattice errors, up to corrections of the order  $\mathcal{O}(a^4, a^2 g_0^{2(n+1)})$  (remember that the Wilson action shows  $\mathcal{O}(a^2)$  deviations already at tree level). That this is possible for all physical quantities at the same time, was proven by Symanzik for  $\varphi^4$  theory and the  $O(3)$   $\sigma$ -model [24], and is also expected to hold for gauge theories [25]. The determination of  $c_i^{(n)}$  is possible by a bootstrap procedure: one determines  $c_i^{(0)}$  by a classical analysis. Then at order  $g_0^2$  physical quantities will show  $\mathcal{O}(a^2)$  deviations, which are to be eliminated by  $c_i^{(1)}$  counterterms; and so forth. Unfortunately lattice perturbation theory is technically cumbersome, especially if larger Wilson loops are included. At one-loop level only one Symanzik-improved action was constructed [26, 27] (previous to our work in chapter 5), and none are known beyond one-loop order.

Due to the renormalization group flow (for  $SU(N)$ )

$$\frac{\partial g_0}{\partial \ln a} = -(\beta_0 g_0^3 + \beta_1 g_0^5 + \mathcal{O}(g_0^7)), \quad \beta_0 = \frac{11N}{48\pi^2}, \quad \beta_1 = \frac{102}{121} \beta_0^2, \quad (1.2.14)$$

the correction terms  $\mathcal{O}(a^2 g_0^{2(n+1)})$  are of the order  $\mathcal{O}(a^2 (\ln a)^{-(n+1)})$ . Therefore the perturbative Symanzik approach cannot be expected to be efficient for all but very small lattice spacings. Indeed, at lattice sizes typically used in Monte Carlo simulations, tree-level and one-loop Symanzik improved actions often show no appreciable improvement over the Wilson action. To overcome this, one should compute  $c_i(g_0^2)$  non-perturbatively, in such a way that the improved action only gives rise to  $\mathcal{O}(a^4)$  corrections. In principle this might be possible to sufficient precision by Monte Carlo techniques. For 'full' QCD, a determination of this kind exists [28] to eliminate the  $\mathcal{O}(a)$  lattice artefacts induced by the fermions. For pure gauge theories at  $\mathcal{O}(a^2)$  the method is complicated by the fact that more Wilson loops are needed for improvement than at  $\mathcal{O}(a)$  in the fermionic sector.

However, an approximate scheme may exist: tadpole improvement [29]. In this scheme, a tree-level<sup>13</sup> Symanzik improved action is modified through a simple prescription, based on mean-field arguments and observations concerning lattice perturbation theory.

The tadpole prescription is to modify the link variables,

$$U_\mu(x) \rightarrow \bar{U}_\mu(x) \equiv U_\mu(x)/u_0, \quad (1.2.15)$$

where  $u_0 \in \mathbb{R}$  is a mean-field parameter, usually defined by

$$u_0 \equiv \left( \frac{1}{N} \langle \text{Re Tr } U_{\mu\nu}(x) \rangle \right)^{\frac{1}{4}}, \quad (\mu \neq \nu). \quad (1.2.16)$$

<sup>12</sup>This terminology refers to the Feynman graph expansion, not to Wilson loops!

<sup>13</sup>Generalization to higher loop order is possible, but usually claimed to be unnecessary.

In this equation,  $\langle \dots \rangle$  is as in eq. (1.2.12), but with  $S_W$  replaced by a tree-level Symanzik improved action<sup>14</sup>, i.e. eq. (1.2.13) for suitable coefficients  $c_i(g_0^2) = c_i$ . Note that  $u_0$  is independent of  $x, \mu, \nu$  due to the lattice symmetries of  $S_{\text{Lat}}$ . However, it will depend on  $g_0$ .

The prescription (1.2.15) can be implemented while still keeping  $U_\mu(x) \in \text{SU}(N)$ , namely by replacing in the path integral

$$S_{\text{Lat}}(\{c_i\}) \rightarrow S_{\text{Lat}}(\{u_0^{4-n_i} c_i\}), \quad (1.2.17)$$

$$g_0^2 \rightarrow \tilde{g}_0^2 u_0^4, \quad (1.2.18)$$

where  $n_i$  is the number of links in a Wilson loop of the class  $C_i$ . Note that eq. (1.2.18) is merely a redefinition of the bare coupling constant (which is why we use a different symbol). This observation shows that tadpole improvement of the Wilson action would have no physical effect.

The latter statement is only valid for a non-perturbative calculation. However, as soon as one makes a finite-order expansion with respect to the coupling constant (such as in the calculation of the Symanzik improvement coefficients  $c_i(g_0^2)$ ), it matters whether  $g_0$  or  $\tilde{g}_0 = g_0/u_0^2$  is used. The reason is that  $u_0^4 = 1 + cg_0^2 + \mathcal{O}(g_0^4)$  with  $c \neq 1$ , so that a finite series in  $\tilde{g}_0$  is a resummation of an infinite series in  $g_0$ . It was already argued by Parisi [30] that  $\tilde{g}_0$  is a more natural expansion parameter than  $g_0$ . His argument was refined by Lepage and Mackenzie [29], who gave many examples for which the applicability of lattice perturbation theory is extended to larger lattice sizes by using a coupling roughly equal to  $\tilde{g}_0$ .

Lepage and Mackenzie also presented a non-rigorous argument why  $\tilde{U}_\mu$  should be considered a more continuum-like variable than  $U_\mu$ . They argued that the most important difference between lattice and continuum gauge theories is the compactness of the link variables  $U_\mu$ . Indeed, in lattice perturbation theory this leads to many tadpole diagrams<sup>15</sup> giving large contributions, while in the continuum formulation tadpoles are completely absent (at least in dimensional regularization). One of the effects is that  $u_0^4 - 1 = cg_0^2 + \mathcal{O}(g_0^4)$ , tending to 0 as  $1/\ln a$ , while from the correspondence (1.2.9) one would expect only  $\mathcal{O}(a^2)$  deviations. It is of course trivial that  $\langle \text{Re Tr } \tilde{U}_{\mu\nu} \rangle \equiv \langle \text{Re Tr } U_{\mu\nu} \rangle_{U \rightarrow \tilde{U}}$  equals 1 exactly, but Lepage and Mackenzie conjectured that the absorption of  $u_0$  into  $U_\mu$  would also improve other lattice operators.

For this reason, the transformation (1.2.15) has been argued (see ref. [31] and references therein) to be applicable to the construction of highly improved lattice gauge theories. Namely, whenever  $S_{\text{Lat}}(\{c_i\})$  is a tree-level Symanzik-improved action,  $S_{\text{Lat}}(\{u_0^{4-n_i} c_i\})$  is conjectured to approximate well an action that is Symanzik improved to all orders in  $g_0$  (potentially even including non-perturbative corrections, if  $u_0$  is measured in a Monte Carlo simulation).

In spite of some attempts [32], even a partial proof of the above line of arguments is unknown. In practice (Monte Carlo), tadpole improvement often is successful

<sup>14</sup>Or rather by the resulting tadpole improved action, eq. (1.2.17), in which case  $u_0$  is to be determined in a self-consistent way.

<sup>15</sup>These are one-loop Feynman (sub)diagrams in which the loop is built from a single propagator.

(e.g. ref. [31]), but less favorable results have also been reported (e.g. ref. [28]). In ref. [31] lattice spacings as big as 0.4 fm were used, rather than 0.05 fm to 0.1 fm for conventional simulations using the Wilson action. Due to this large potential gain, a method to predict, for a given quantity, the reliability of a tadpole-improved Monte Carlo simulation would be extremely welcome.

### 1.3 Periodic boundary conditions and twist

This section mainly aims to give some physical background to chapter 5, which is a rather technical chapter. We discuss the concept of twisted boundary conditions in  $SU(N)$  pure gauge theories, introduced by 't Hooft [10]. For convenience we restrict ourselves to the continuum formulation. However, twist is also well defined on the lattice [33], so that the analysis below can be brought over to the formulation needed in chapter 5. This is particularly straightforward if one introduces lattice twist as in ref. [27].

't Hooft's analysis starts with the observation that choosing periodic boundary conditions for physical (i.e. gauge invariant) fields, still allows for generalized boundary conditions on  $A_\mu$ :

$$A_\mu(x_\nu = L_\nu) = A_\mu^{\Omega_\nu}(x_\nu = 0), \quad (1.3.1)$$

where the other coordinates  $x_{\lambda \neq \nu}$  are free.  $L_\nu$  denotes the size of the box in the  $\nu$  direction, and  $A_\mu^{\Omega_\nu}$  is a gauge transformation, defined in eq. (1.1.2). The so-called twist matrix  $\Omega_\nu$  is independent of  $x_\nu$ , but may depend on the other coordinates.

A subtle structural requirement is:

$$\Omega_\mu(x_\nu = L_\nu)\Omega_\nu(x_\mu = 0) = \Omega_\nu(x_\mu = L_\mu)\Omega_\mu(x_\nu = 0)z^{n_{\mu\nu}}, \quad (z \equiv e^{2\pi i/N}), \quad (1.3.2)$$

for each  $\mu, \nu$  plane. This is a consistency relation, necessary because there are two ways to relate  $x$  to  $x + L_\mu \hat{\mu} + L_\nu \hat{\nu}$ .  $n_{\mu\nu} = -n_{\nu\mu}$  are integers, only defined modulo  $N$  (i.e.  $n_{\mu\nu} \in \mathbb{Z}_N$ ). This guarantees that  $Z_{\mu\nu} \equiv \exp(2\pi i n_{\mu\nu}/N)$  is a center element of  $SU(N)$ , i.e.:  $Z_{\mu\nu} \in SU(N)$  and for all  $U \in SU(N)$ ,  $[Z_{\mu\nu}, U] = 0$ . In *pure* gauge theories it is not required that  $Z_{\mu\nu} = 1$ , because  $A_\mu(x)$  is invariant under gauge transformations with a center element.

To a large extent the twist matrices are unphysical objects. The reason is that, generically, a gauge transform  $A_\mu^\Omega$  of  $A_\mu$  will not satisfy eq. (1.3.1) ( $\Omega$  is not restricted by any boundary conditions). Instead,  $A_\mu^\Omega$  satisfies twisted boundary conditions with gauge transformed twist matrices:

$$\Omega_\mu \rightarrow \Omega_\mu^\Omega \equiv \Omega(x_\mu = L_\mu)\Omega_\mu\Omega^{-1}(x_\mu = 0). \quad (1.3.3)$$

However, for any  $\Omega$  the set of  $\Omega_\mu^\Omega$  satisfies eq. (1.3.2) with the same  $n_{\mu\nu}$ . This means that the twist tensor  $n_{\mu\nu}$  cannot be gauge transformed away. In fact, it is the only gauge invariant information carried by the twist matrices<sup>16</sup>. If  $n_{0i} = 0 \pmod{N}$ , it is

<sup>16</sup>In this subsection we neglect instanton effects [19]. See ref. [34] for a detailed account of topology on a four-torus.

possible to render the twist matrices  $\Omega_i$   $x$ -independent by a gauge transformation. Below we will assume that this has been done.

How pure gauge theory is quantized in the presence of twist, is described in refs. [10, 34]. For convenience, we give a very short summary of the arguments and the results. In a Hamiltonian approach, one first drops the dimension of time. The theory can then be quantized separately for each spatial boundary condition implied by  $\Omega_i$ . Physical states are required to be invariant under gauge transformations  $\Omega$  respecting the boundary conditions, i.e.  $\Omega \in B_{\{(\mathbf{n}, \mathbf{j})\}}$  with

$$B_{\{(\mathbf{n}, \mathbf{j})\}} \equiv \{\Omega \mid \forall i \Omega(x_i = L_i) = \Omega_i \Omega(x_i = 0) \Omega_i^{-1}\}. \quad (1.3.4)$$

(A gauge transformation  $|\psi\rangle^\Omega$  of a state  $|\psi\rangle$ , also denoted by  $\Omega|\psi\rangle$ , is defined by  $\langle A|\psi\rangle^\Omega \equiv \langle A^{\Omega^{-1}}|\psi\rangle$  for all 'coordinates' (spatial vector potentials)  $A$ ). However, again because we are considering a *pure* gauge theory, there is a wider class of gauge transformations that commute with the Hamiltonian ( $n_j \in \mathbb{Z}_N$ ):

$$B_{\{(\mathbf{n}, \mathbf{j})\}}(\mathbf{n}) \equiv \{\Omega \mid \forall i \Omega(x_i = L_i) = z^{n_i} \Omega_i \Omega(x_i = 0) \Omega_i^{-1}\}. \quad (1.3.5)$$

For  $\mathbf{n} \neq 0 \pmod{N}$ , physical states are *not* required to be invariant under  $\Omega[\mathbf{n}] \in B_{\{(\mathbf{n}, \mathbf{j})\}}(\mathbf{n})$ . We can therefore split the Hilbert space in eigenspaces of  $\Omega[\mathbf{n}]$  (it does not matter which representative of  $B_{\{(\mathbf{n}, \mathbf{j})\}}(\mathbf{n})$  is chosen, because different elements of  $B_{\{(\mathbf{n}, \mathbf{j})\}}(\mathbf{n})$  are related by multiplication with an element of  $B_{\{(\mathbf{n}, \mathbf{j})\}}(0) = B_{\{(\mathbf{n}, \mathbf{j})\}}$ ).

To this end let us define the set of  $N^3$  distinct projection operators ( $e_i \in \mathbb{Z}_N$ )

$$P[\mathbf{e}] \equiv \frac{1}{N^3} \sum_{\mathbf{n} \in \mathbb{Z}_N^3} z^{-\mathbf{n} \cdot \mathbf{e}} \Omega[\mathbf{n}]. \quad (1.3.6)$$

Since  $\Omega[\mathbf{n}]\Omega[\mathbf{n}'] \in B_{\{(\mathbf{n}, \mathbf{j})\}}(\mathbf{n} + \mathbf{n}')$ , it directly follows that (for a physical state  $|\psi\rangle$ )

$$\Omega[\mathbf{n}]P[\mathbf{e}]|\psi\rangle = z^{\mathbf{n} \cdot \mathbf{e}} P[\mathbf{e}]|\psi\rangle. \quad (1.3.7)$$

The quantum theory thus has  $N^3$  distinct sectors, in which all  $\Omega[\mathbf{n}]$  are diagonalized. The Hamiltonian can be diagonalized separately in each sector.

In conclusion, twisted boundary conditions allow for  $N^6$  quantum sectors:  $N^3$  due to (physically) periodic boundary conditions in space, labeled by  $n_{ij}$ , and within each of these:  $N^3$  sectors labeled by  $\mathbf{e}$ . It should be noted that the latter can also be found in a path integral approach, because from eqs. (1.3.2), (1.3.5) it is clear that  $\Omega[\mathbf{n}]$  can be identified with a twist matrix  $\Omega_0$ .

These sectors have a beautiful interpretation, due to 't Hooft [10]:  $\mathbf{e}$  corresponds to electric flux,  $\mathbf{m}$  to magnetic flux (with  $n_{ij} = -\sum_k \epsilon_{ijk} m_k$ ). For shortness we will not repeat 't Hooft's arguments (see also ref. [34]) why this is a natural identification. Let us only mention that a gauge invariant (under transformations in the class  $B_{\{(\mathbf{n}, \mathbf{j})\}}(0)$ ) operator that takes the  $\mathbf{e}^{\text{th}}$  electric sector into the  $(\mathbf{e} + \hat{\mathbf{j}})^{\text{th}}$  sector, is the Polyakov line  $P_j$ :

$$P_j(\mathbf{x}^{(j)}) \equiv \text{Tr} \left[ P \exp \left( \int_0^{L_j} dx_j A_j(\mathbf{x}) \right) \Omega_j \right] \quad (1.3.8)$$

(where  $\mathbf{x}^{(1)} \equiv (x_2, x_3)$ , and cyclic). This can be deduced from eq. (1.2.4) and the definition of  $\Omega[\mathbf{n}]$ .

We would now like to apply the above results to the specific geometry used in chapter 5. To this end we restrict ourselves to a magnetic flux  $\mathbf{m} = (0, 0, -1)$ . We also choose  $L_1 = L_2 \equiv L$  and send  $L_3 \rightarrow \infty$ , thus obtaining a 'twisted tube'. (This terminology is due to Lüscher and Weisz, whose work [26, 27] inspired us to do the calculation in chapter 5). It can be proven that the vacuum valley of the twisted tube consists of only one point,  $A_\mu = 0$  (up to a gauge transformation), which is why it admits a straightforward perturbative analysis, contrary to the purely periodic case  $\mathbf{m} = 0$  (cf. eq. (1.1.4)). Remember that perturbation theory is applicable for small values of  $L$ , due to asymptotic freedom.

What we want to do in the remainder of this section, is to classify the lowest-lying excitations above the quantum vacuum  $|0\rangle$ . To this end we use a path integral approach. We can get a handle on the spectrum by considering the *gauge invariant* correlation function of two Polyakov lines at a distance  $t$  in (Euclidean) time,

$$\langle P_1^\dagger(x_2, x_3; t) P_1(x_2, x_3; 0) \rangle. \quad (1.3.9)$$

Here  $\langle \dots \rangle$  is the continuum analogue of eq. (1.2.12), with integration over fields that satisfy twisted boundary conditions (determined by  $\Omega_{1,2}$ ) in  $x_{1,2}$  and arbitrary boundary conditions in  $x_{0,3}$ . Since  $|0\rangle$  is a state with zero electric flux, this correlation function corresponds to the creation of a (gauge invariant, i.e. physical) state at time 0, with electric flux  $\mathbf{e} = (1, 0, 0)$ , and its annihilation at time  $t$ . Therefore for large values of  $t$  it will be dominated by the exponential behavior  $\exp(-Mt)$ , where  $M$  is the energy of the lowest-lying Hamiltonian eigenstate that has non-zero overlap with  $\hat{P}_1(x_2, x_3)|0\rangle$ .

In order to compute eq. (1.3.9) perturbatively, the allowed fields in the path integral must be parametrized. To this end one defines the  $\mathbf{x}$ -independent matrices

$$\Gamma[\mathbf{n}] = \Omega_1^{-n_2} \Omega_2^{n_1} z^{\frac{1}{2}(n_1+n_2)(n_1+n_2-1)}, \quad \mathbf{n} \in \mathbb{Z}_N^3. \quad (1.3.10)$$

The choice of the phase factor is a matter of convenience. The important property of  $\Gamma[\mathbf{n}]$  is that it is an element of  $B_{\{(\Omega_i)\}}(-\mathbf{n})$ . From the second reference in [33] we know that

$$\frac{1}{N} \text{Tr} \left( \Gamma[\mathbf{n}] \Gamma^\dagger[\mathbf{n}'] \right) = \delta_{\mathbf{n}, \mathbf{n}' \pmod{N}}. \quad (1.3.11)$$

This allows us to Fourier transform  $A_\mu(x)$  in the following way:

$$A_\mu(x) = \frac{g_0}{(2\pi)^2 L^2 N} \sum_{k_1, k_2} \int_{\mathbb{R}^2} dk_0 dk_3 e^{ikx} \bar{A}_\mu(k) \Gamma_k. \quad (1.3.12)$$

Here  $kx \equiv \sum_{\mu=0}^3 k_\mu x_\mu$  and  $\sum_{k_1, k_2}$  runs over

$$k_\nu = m n_\nu, \quad m \equiv \frac{2\pi}{NL}, \quad n_\nu \in \mathbb{Z}, \quad (\nu = 1, 2). \quad (1.3.13)$$

Also we adopted the notation of ref. [27]:

$$\Gamma_{\mathbf{k}} \equiv \Gamma[(n_1, n_2, 0)], \quad (k_\nu = mn_\nu, \quad \nu = 1, 2). \quad (1.3.14)$$

In contrast to the purely periodic case,  $\bar{A}_\mu(k) \in \mathbb{C}$  carries no  $SU(N)$  index. Instead, the momenta in the twisted directions are quantized in units of  $m$  rather than  $2\pi/L = Nm$ . The total number of degrees of freedom is preserved, because  $\text{Tr} A_\mu(x) = 0$  implies

$$\bar{A}_\mu(k) = 0 \quad \text{for } n_{1,2} = 0 \pmod{N}. \quad (1.3.15)$$

Using this formalism, one quickly derives (with  $t = x_0$ )

$$P_1(x_2, x_3; t) = \frac{g_0}{(2\pi)^2 L} \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^2} dk_0 dk_3 e^{ikx} \bar{A}_1(k) + \mathcal{O}(A_1^2), \quad (1.3.16)$$

$$k = (k_0, 0, (2\pi/L)n + m, k_3). \quad (1.3.17)$$

Note that  $k_2$  is now quantized in units of  $Nm$ , but carries an additional momentum  $m$ . Physically, this is the Poynting vector  $\mathbf{e} \times \mathbf{m}$  due to the electric flux  $\mathbf{e} = (1, 0, 0)$  created by the Polyakov line [35].

For the moment we neglect the  $\mathcal{O}(A_1^2)$  corrections in eq. (1.3.16). We then obtain

$$\langle P_1^\dagger(x_2, x_3; t) P_1(x_2, x_3; 0) \rangle = \frac{g_0^2 N}{8\pi^2} \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^2} dk_0 dk_3 e^{ik_0 t} D_{11}^{\text{full}}(k), \quad (1.3.18)$$

where  $D_{11}^{\text{full}}(k)$  is the dressed propagator, defined through the 2-point function:

$$\langle \bar{A}_\mu^*(k) \bar{A}_\nu(k') \rangle = \frac{1}{2} ((2\pi)^2 L^2 N) \delta_{n'_1, m_1} \delta_{n'_2, n_2} \delta(k'_0 - k_0) \delta(k'_3 - k_3) \chi_k D_{\mu\nu}^{\text{full}}(k). \quad (1.3.19)$$

The factor  $\chi_k \equiv (1 - \delta_{n_1, 0 \pmod{N}} \delta_{n_2, 0 \pmod{N}})$  is implied by eq. (1.3.15): modes with zero momentum ( $\pmod{Nm}$ ) in the twisted directions are absent on the twisted tube. In eq. (1.3.19),  $k_\nu = mn_\nu$  and  $k'_\nu = mn'_\nu$  ( $\nu = 1, 2$ ). In eq. (1.3.18), the momentum  $k$  is restricted to eq. (1.3.17).

Now it is well known that

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}} dk_0 e^{ik_0 t} D_{11}^{\text{full}}(k) \sim e^{-E(\mathbf{k})t}, \quad (1.3.20)$$

where  $k = (iE(\mathbf{k}); \mathbf{k})$  is the pole of  $D_{11}^{\text{full}}(k)$ . To lowest order in perturbation theory this gives the usual formula  $E(\mathbf{k}) = |\mathbf{k}| \equiv \sqrt{k_1^2 + k_2^2 + k_3^2}$ . Here it is essential that the Polyakov line, due to its gauge invariance, couples only to physical (i.e. transversal) polarizations. Indeed, from eq. (1.3.16) we see that  $P_1(x_2, x_3)$  couples to the polarization  $\varepsilon_\mu = \delta_{\mu, 1}$ , which is transversal due to eq. (1.3.17).

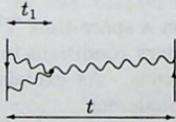
Inserting this result in eq. (1.3.18), we see that the leading exponential decay comes from  $\mathbf{k} = (0, m, 0)$ . Thus the lightest particle with electric flux  $(1, 0, 0)$  has a mass  $m$  ( $+\mathcal{O}(g_0^2)$ ). This is what Lüscher and Weisz called an A meson. From eqs. (1.3.13), (1.3.15) we see that it is the lightest particle on the twisted tube.

$(e_1, e_2)$	$CP^{(1)}(e_1, e_2)$	$CP^{(2)}(e_1, e_2)$
(1, 0)	(1, 0)	(0, -1)
(0, 1)	(0, -1)	(-1, 0)
(-1, 0)	(-1, 0)	(0, 1)
(0, -1)	(0, 1)	(1, 0)

Table 1-1. Properties of the electric fluxes created by the Polyakov lines  $P_1$ ,  $P_2$ ,  $P_1^I$  and  $P_2^I$ , under two CP transformations:  $P^{(1)}(x_1, x_2) = (-x_1, x_2)$  and  $P^{(2)}(x_1, x_2) = (x_2, x_1)$ .

There are more particles with the same mass. To lowest order in perturbation theory these are associated with the other poles of the bare propagator for  $|\mathbf{k}| = m$ , i.e. at  $\mathbf{k} = (m, 0, 0)$ ;  $(0, -m, 0)$ ;  $(-m, 0, 0)$ . Of course they are created by Polyakov lines with electric fluxes  $\mathbf{e} = (0, -1, 0)$ ;  $(-1, 0, 0)$ ;  $(0, 1, 0)$ . It can be proven that the total of four particles now found have equal masses to all orders in  $g_0$ . This is simply a consequence of symmetry considerations. Namely, the Hamiltonian has two CP symmetries, one corresponding to (1)  $x_1 \rightarrow -x_1$ , and the other corresponding to (2)  $x_1 \leftrightarrow x_2$ . We will not prove this here (for the proof see ref. [27]). Let us only mention that (for  $N \geq 3$ ) the charge conjugation C (defined by  $A_\mu(x) \rightarrow W A_\mu^*(x) W^{-1}$ , where  $W \in SU(N)$  can be freely chosen) is not a symmetry, because it violates the boundary conditions (it changes  $\mathbf{m}$  into  $-\mathbf{m}$ ). The behavior of the electric fluxes under the two CP transformations (referred to as  $CP^{(1,2)}$ ) is given in table 1-1.

One may wonder about the  $\mathcal{O}(A_\mu^2)$  corrections in eq. (1.3.17). In principle they bring about  $\mathcal{O}(g_0^2)$  corrections to the above analysis, because they give rise to 3-point (or in general  $n$ -point,  $n \geq 3$ ) functions in the expansion of (1.3.9). The leading Feynman diagram of this type is given by (in coordinate space)



$$(1.3.21)$$

For  $t \rightarrow \infty$ , we must have  $t_1/t \rightarrow 0$ . The reason is that from the above analysis we know that each propagator of length  $t'$  suppresses the correlation function by a factor of approximately  $\exp(-mt')$ . Thus only diagrams in which a single (dressed) propagator runs from the one Polyakov line to the other, contribute to the leading behavior for  $t \rightarrow \infty$ :  $\exp(-(m + \mathcal{O}(g_0^2))t)$ . Diagrams like the one above do not contribute to the  $\mathcal{O}(g_0^2)$  corrections to the mass  $m$ . Instead, their contributions should be associated with the renormalization of the composite operator  $P_1(x_2, x_3)$ .

There is a further set of four particles which at tree-level are degenerate with the four particles found above. These also carry electric flux  $|\mathbf{e}| = 1$ , but they have a different polarization,  $\varepsilon_\mu = \delta_{\mu,3}$ , at momenta  $\mathbf{k} = (0, m, 0)$ ;  $(m, 0, 0)$ ;  $(0, -m, 0)$ ;  $(-m, 0, 0)$ . We will refer to this set as  $A^-$  mesons, because under appropriate CP

transformations they have eigenvalues opposite to their counterparts found above. Those we will call  $A^+$  mesons. Due to the  $CP^{(1,2)}$  symmetries, all  $A^-$  mesons are exactly degenerate. However, there is no symmetry relating the  $A^-$  mesons to the  $A^+$  mesons, so beyond tree-level they must be expected to have different masses. The  $A^-$  mesons can be created by gauge-invariant operators, but these are of a more complicated nature than the Polyakov lines creating the  $A^+$  mesons.

No more particles with masses  $m + \mathcal{O}(g_0^2)$  exist, because the Lagrangian (1.1.1) admits only two physical polarizations. From the bare propagator one sees that the next particles<sup>17</sup> encountered in the spectrum have twisted momenta  $|k_1| = |k_2| = m$ , and thus their energies equal  $\sqrt{2}m + \mathcal{O}(g_0^2)$ . These were called B mesons by Lüscher and Weisz. In our language they carry electric flux  $|e_1| = |e_2| = 1$ . Like for  $A^+$  mesons, one of the two possible polarizations of B mesons is created by simple operators, for example

$$P_{12}(x_3) \equiv \text{Tr} \left[ P \exp \left( \int_0^L dx_1 A_1(x_1, 0, x_3) \right) P \exp \left( \int_0^L dx_2 A_2(L, x_2, x_3) \right) \Omega_1 \Omega_2 \right]. \quad (1.3.22)$$

Still higher up are particles with masses  $2m + \mathcal{O}(g_0^2)$  or more. For  $g_0 \neq 0$  many of these particles should be expected to be unstable, because the kinematics allow them to decay in A or B mesons (only the particles with tree-level masses  $2m$  might be safe, depending on the  $\mathcal{O}(g_0^2)$  corrections).

## 1.4 Outline

In this section we explain how the topics discussed above come together in the remainder of this thesis.

In chapter 2 we perform a search for  $SU(2)$  instantons on a space-time  $T^3 \times \mathbb{R}$ , with periodic<sup>18</sup> boundary conditions in space, and free boundary conditions in time. Our motivation for this is the possibility of finding sphalerons. As explained in section 1.1, we expect the widest instantons to go through sphalerons.

Also lattice gauge theory comes into play, because instantons will be looked for by numerical methods. One might suspect that we will be using a tree-level improved lattice action, but in fact an 'over-improved' action is what we need. For this action the leading lattice artefacts have opposite sign compared to the Wilson action, and this allows us to obtain stable numerical results for the largest instantons.

In chapter 2 we are also interested in the question how the non-zero dimensionality of the vacuum valley relates to the dimension of the instanton moduli-space. However, due to the numerical nature of our analysis, the question cannot be completely

<sup>17</sup>It is natural to consider particles with momenta  $(0, m, k_3)$ ,  $k_3 \neq 0$ , as A mesons in motion.

<sup>18</sup>For shortness, purely periodic boundary conditions, i.e. periodicity of  $A_\mu(x)$ , will usually be referred to as periodic boundary conditions. The other physically periodic boundary conditions discussed in the previous section, will be called twisted boundary conditions.

resolved. This is why we make a little side step in chapter 3, where we consider the  $O(3)$   $\sigma$ -model on a two-dimensional space-time  $T^1 \times \mathbb{R}$ . This model is in some important respects much alike  $SU(2)$  pure gauge theory on  $T^3 \times \mathbb{R}$ . In particular it also possesses a non-trivial vacuum valley. For the  $O(3)$   $\sigma$ -model we will be able to find all instantons analytically, and also all sphalerons. Thus we can check the conjectures made in chapter 2 in the context of the  $O(3)$   $\sigma$ -model.

In chapters 4 and 5 we turn our attention to the improvement of lattice actions. In particular in chapter 4 we put tadpole improvement to a critical examination. We will do so, once more, for  $SU(2)$  pure gauge theory in a (purely periodic) spatial volume  $T^3$ . The reason is that for this geometry the low-lying spectrum has been obtained analytically in the continuum [8] as well as for the Wilson lattice action [11], up to volumes where some non-perturbative effects are clearly visible. We will be able to partly repeat the calculation for a judiciously chosen tree-level Symanzik improved action, called the square action. Also we will present results of Monte Carlo simulations for various lattice actions.

Chapter 5 is concerned with the computation of the one-loop corrections to the Symanzik improvement coefficients  $c_i$  (eq. (1.2.13)) for the square action. Before the writing of this thesis, such a calculation had only been done [26, 27] for the Lüscher-Weisz Symanzik-improved action. We consider it useful to have at least one alternative around, because it is potentially dangerous to draw conclusions concerning improvement based on a single choice of action.

Since the improvement coefficients should not depend on the details of the theory at physical distances, one is free to use boundary conditions that are best suited for a perturbative calculation. We will use this freedom to calculate the coefficients on the twisted tube. As a consistency check, on the space-time  $\mathbb{R}^4$  we will compute the potential between two static quarks at small distances. We end chapter 5 by comparing our results to the tadpole prediction.

In the final chapter we put the results of the preceding chapters in perspective.



## 2 Instantons from over-improved cooling

### 2.1 Introduction

Since the time of the discovery of instantons [17] in non-Abelian gauge theories, as vacuum to vacuum quantum mechanical tunneling events [19], their role in strongly interacting theories has been controversial, both in the continuum [36] and in the lattice formulation [37]. For the continuum this has been mainly due to applying semiclassical techniques, which cannot be justified at strong coupling. In the lattice formulation the main problems were the instantons localized at the scale of the lattice cut-off for which topological charge cannot be defined unambiguously [38], and which have actions considerably lower than the continuum action of  $8\pi^2$ . Strictly speaking, there are no locally stable solutions on a lattice using the standard Wilson action [7], because this lattice action decreases when the instanton becomes more localized [39], as we will demonstrate also from analytic considerations. On a trial and error basis, different (improved) lattice actions were considered, some of them indeed giving rise to stable lattice solutions [40]. This chapter will provide the proper framework to understand the stability.

It is not too difficult to see the reason for (in)stability. At finite lattice sizes the lattice action deviates from the continuum and this deviation is larger for stronger fields. For the Wilson action, as we will show, the lattice artefacts make the action decrease as compared to the continuum. In the continuum, instantons have a scale (or size) parameter  $\rho$ , on which the action does not depend. But the smaller  $\rho$  becomes, the larger the fields get, which makes the lattice action decrease. On dimensional grounds one easily argues that (generically)  $S_{\text{Lat}}(a, \rho) = 8\pi^2(1 + (a/\rho)^2 d_2 + \mathcal{O}(a/\rho)^4)$  for  $\rho \gg a$ , which will be demonstrated in more detail further on. For the Wilson action  $d_2 < 0$ , explaining the instability. Hence one simply modifies the action such that  $d_2 > 0$ , in order to obtain stable solutions for the maximal value of  $\rho$  allowed by the volume  $[0, L]^3$ , which is kept finite. The modified action need not be of the type of an improved action [24, 25], for which typically one wants to achieve  $d_2 = 0$ . In that case, as we will show, the  $(a/\rho)^4$  term might still destabilize the solution.

We deliberately want to keep  $d_2 > 0$ , which we will hence call over-improvement. The reason is that our motivation for embarking on this project was to find the instantons with the largest scale  $\rho$ . These presumably correspond to tunneling over the lowest energy barrier, separating two classical vacua. The configuration that corresponds to the lowest barrier height is then conjectured to be a sphaleron (which exists due to the fact that we keep the volume finite). Remember that a sphaleron [13] is a saddle point of the energy functional (i.e. the potential) with precisely one unstable direction. If indeed the sphaleron is on the top of an instanton path, then this path reaches the sphaleron along the unstable direction.

Assuming the conjecture to be true, we can use the instantons to map out that part of configuration space where, at some point when increasing the volume  $[0, L]^3$ , quantum mechanical wave functions will show a sudden increase in amplitude. We refer to a study on  $S^3 \times \mathbb{R}$  [15, 16] for readers interested in this issue. In ref. [16] it was also argued that extension of the calculation to the geometry  $T^3 \times \mathbb{R}$  would be profitable. There are two reasons for this, both related to the rectangular nature of  $T^3 \times \mathbb{R}$ . First,  $T^3 \times \mathbb{R}$  can be discretized to a cubic lattice, and thus analytic results can be compared to Monte Carlo simulations. Second,  $\mathbb{R}^3$ , i.e. the spatial volume that is physically relevant, can be partitioned in cubes, but not in spheres. This is important if one wants to analyze whether the quantum vacuum of gauge theory on  $\mathbb{R}^3$  exhibits domain formation, as often has been speculated [41].

In this chapter we will study instantons on  $T^3 \times \mathbb{R}$ . This geometry allows us to find the instantons using the lattice approximation. Though the main goal of this chapter is to analyze the  $\rho$  dependence introduced by the discretization, we will also pursue knowledge of the instanton moduli-space on  $T^3 \times \mathbb{R}$ . It should be understood that both issues are of a purely classical nature.

The outline of this chapter is as follows. In section 2.2 we present known facts about continuum instanton solutions. In section 2.3 we describe the well-known cooling algorithm [39, 42] that is used to find minima of lattice actions. In section 2.4 we systematically expand the Wilson action (and other lattice actions) around the continuum action to compute the coefficient  $d_2$  introduced above, and in section 2.5 we compute additional corrections that sometimes must be taken into account. Numerical results obtained by our over-improvement method are discussed in section 2.6. For simplicity we restrict ourselves to  $SU(2)$  pure gauge theories.

## 2.2 On the existence of continuum solutions

The geometry  $T^3 \times \mathbb{R}$ , in particular in a lattice formulation, can be seen as a limiting case of an asymmetric four-torus  $[0, L]^3 \times [0, T]$ . The only known solutions have constant curvature [43] and hence cannot correspond to vacuum to vacuum tunneling. Furthermore their topological charge is at least 2. Actually, it can be proven rigorously [44] that for  $T$  finite, no regular charge 1 self-dual solutions can exist on a four-torus (we will illustrate this with our numerical results). As soon as we allow for twisted boundary conditions [10], existence of instanton solutions with minimal non-trivial topological charge can be proven. One distinguishes two cases, depending on the properties of the twist tensor  $n_{\mu\nu} \in \mathbb{Z}_2$ .

If  $\frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}n_{\mu\nu}n_{\lambda\sigma} = 1 \pmod{2}$  (called non-orthogonal twist), the topological charge is half-integer. The minimal action allowed by the topological bound is therefore  $4\pi^2$ , corresponding to topological charge  $1/2$ . As twist is also well defined on the lattice [33], and for non-orthogonal twist does not allow for zero-action configurations, these instantons cannot 'fall through the lattice'<sup>1</sup>. Indeed, the index theorem [45]

<sup>1</sup>This terminology refers to what usually is observed if cooling has decreased the instanton size

predicts in this case 4 parameters ( $8 \times$  topological charge), which have to correspond to the position parameters. The charge  $1/2$  instanton hence has fixed size and cannot shrink due to lattice artefacts. Impressively accurate results [46] were obtained for this case using the cooling method [39,42] to find a solution to the (lattice) equations of motion. Its smoothness and scaling with the lattice volume leaves no room to doubt it provides an accurate approximation to the continuum solution with action  $4\pi^2$ . In the continuum, existence of smooth non-trivial (but not necessarily self-dual) solutions was proven by Sedlacek [47], whereas theorem 3.2.1. of ref. [48] states that the moduli space of self-dual solutions with topological charge  $1/2$  is isomorphic to a four-torus.

If  $\frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}n_{\mu\nu}n_{\lambda\sigma} = 0 \pmod{2}$  (called orthogonal twist), there are 'twist eating' [33] configurations, i.e. configurations that have zero action and are compatible with twisted boundary conditions (see also ref. [49]). For  $SU(2)$  it was shown in ref. [50] that as long as  $n_{\mu\nu} \neq 0 \pmod{2}$  for some  $\mu$  and  $\nu$ , this twist eating configuration is unique, up to a global gauge transformation if a twist is introduced as in ref. [33,46], and up to multiplication with elements of the center of the gauge group. With twisted boundary conditions as originally defined by 't Hooft [10] a global gauge transformation would even change the boundary conditions, and as  $SO(3)$  bundle the twist eating configuration is unique (For  $SU(N)$  it can be proven [51] that out of the  $N^4$  center elements that can multiply the twist, only  $N^2$  give rise to gauge-inequivalent configurations). Under this condition it can be shown [48] that there are instanton solutions with 8 parameters (its moduli space, when dividing out the trivial translation parameters, is even related to a  $K3$  surface [48,52]). This is done by using Taubes' [53] technique of gluing a localized instanton (with scale, position and global gauge parameters) to the twist eating flat connection (i.e. zero-action configuration). As the latter is not invariant under global gauge transformations, the global gauge parameters of the localized instantons are genuine parameters of the moduli space (see also ref. [54]).

The reason twisted boundary conditions are useful, is that at finite  $T$  there are no exact instantons on  $T^4$  with periodic boundary conditions, but there are exact solutions for any non-trivial twist in the time direction. As  $T \rightarrow \infty$  these solutions are also solutions on  $T^3 \times \mathbb{R}$ . This comes about as follows. Since at  $T \rightarrow \infty$  the action can only stay finite if  $|t| \rightarrow \infty$  the energy density goes to zero, we deduce from a vanishing magnetic energy that up to a gauge

$$A_j(\mathbf{x}, t \rightarrow \pm\infty) = \frac{1}{2}i\sigma_3 \frac{C_j^\pm}{L}, \quad (2.2.1)$$

where  $C_j^\pm \in [0, 4\pi]$  ( $A_0 = 0$ ) parametrizes the vacuum or toron valley [8], whose gauge invariant observables are best described by the Polyakov line expectation values

$$P_i \equiv \frac{1}{2} \text{Tr} \left\{ P \exp \left( \int_0^L dx_i A_i(\mathbf{x}, t) \right) \right\} = \cos(C_i/2). \quad (2.2.2)$$

to  $\rho \approx a$ : after some additional cooling the action suddenly drops to zero.

(For the proper definition in the presence of twist, see eq. (1.3.8). Also note that in the present chapter we choose to include a normalization factor). In the vacuum valley,  $P_i$  is space independent and the vanishing of the electric energy at  $t \rightarrow \pm\infty$  also requires  $P_i$  (or  $C_i^\pm$ ) to be asymptotically time independent. Instanton solutions on  $T^3 \times \mathbb{R}$  are hence (partly) characterized by the boundary conditions  $C_i^\pm$  at  $t \rightarrow \pm\infty$ . It is these general instantons that are physically relevant.

It is not clear if solutions exist with arbitrary boundary values. Approaching  $T \rightarrow \infty$  by using periodic boundary conditions (i.e.  $A_\mu(\mathbf{x}, t \rightarrow \infty) = A_\mu(\mathbf{x}, t \rightarrow -\infty)$ ) up to a periodic gauge transformation, imposing  $C_i^+ = C_i^- \bmod 4\pi$  does not allow us to prove existence. As long as  $T$  is finite there are no solutions [44] and the proof of non-existence breaks down as  $T \rightarrow \infty$ . On the other hand, with twist in the time direction<sup>2</sup>,  $n_{0i} = 1$ , even at  $T$  finite there is in the continuum an 8 parameter set of exact instanton solutions, which due to the twisted boundary conditions are constrained to  $P_i(t = 0) = (-1)^{n_{0i}} P_i(t = T)$ . For  $T \rightarrow \infty$  this implies  $C_i^+ = (2\pi - C_i^-) \bmod 4\pi$ . For localized instantons, asymptotically the field has to coincide with the unique flat connection, which fixes the possible values of  $C_i^\pm$  to  $\pi$ , but at the other extreme, as the instanton in the spatial direction extends up to the 'boundary' of the torus, the regions  $t \rightarrow +\infty$  and  $t \rightarrow -\infty$  no longer are connected, which will relax the condition  $P_i = 0$  ( $C_i^\pm = \pi$ ). Although we have no proof, it is reasonable to assume that the 8 parameters for the instantons close to the maximal size are described by  $\rho$ , the 4 position parameters and the 3 vacuum valley parameters  $C_i^+$  (or  $C_i^-$ ). Note that for  $P_i \rightarrow 0$  as  $t \rightarrow \pm\infty$ , the solution at infinite  $T$  is compatible with twisted *as well as* periodic boundary conditions. In any case we have now learned that on  $T^3 \times \mathbb{R}$  (i.e. with free boundary conditions at  $t \rightarrow \pm\infty$ ) there are at least 8 and at most 11 continuous parameters that describe the instanton solutions for vacuum to vacuum tunneling.

## 2.3 The lattice actions and cooling

Let us start with discussing the standard Wilson action [7]

$$S_W = \sum_{x, \mu, \nu} \text{Tr} \left( 1 - \nu \prod_{x \xrightarrow{\mu}} 1 \right) = \sum_{x, \mu, \nu} \text{Tr} \left( 1 - U_\mu(x) U_\nu(x + a\hat{\mu}) U_\mu^\dagger(x + a\hat{\nu}) U_\nu^\dagger(x) \right), \quad (2.3.1)$$

where  $U_\mu(x)$  are  $SU(2)$  group elements on the link that runs from  $x$  to  $x + a\hat{\mu}$ ,  $\hat{\mu}$  being the unit vector in the  $\mu$  direction. To derive the equations of motion, we observe that  $S_W$  depends on  $U_\mu(x)$  through the expression:

$$S_W(U_\mu(x)) = \text{Tr} \left( 1 - U_\mu(x) \bar{U}_\mu^\dagger(x) \right) + \text{Tr} \left( 1 - U_\mu^\dagger(x) \bar{U}_\mu(x) \right), \quad (2.3.2)$$

<sup>2</sup> $n_{0i} = 1$  can be implemented with periodic boundary conditions in space by  $A_\mu(\mathbf{x}, T) = \Omega_0(A_\mu(\mathbf{x}, 0) + \hat{\partial}_\mu) \Omega_0^{-1}$ , with  $\Omega_0 = \exp(\pi i \sum_j n_{0j} x_j \sigma_j / L)$ , which is an antiperiodic Abelian gauge transformation.

where

$$\begin{aligned} \tilde{U}_\mu(x) = & \sum_{\nu \neq \mu} \left( \nu \begin{array}{c} \mu \\ \downarrow \\ \mu \end{array} + \nu \begin{array}{c} \mu \\ \leftarrow \\ \mu \end{array} \right) = \sum_{\nu \neq \mu} \left( U_\nu(x) U_\mu(x + a\hat{\nu}) U_\nu^\dagger(x + a\hat{\mu}) \right. \\ & \left. + U_\nu^\dagger(x - a\hat{\nu}) U_\mu(x - a\hat{\nu}) U_\nu(x + a\hat{\mu} - a\hat{\nu}) \right), \end{aligned} \quad (2.3.3)$$

which is independent of  $U_\mu(x)$ . Hence the principle of 'least' action,  $S_W(e^X U_\mu(x)) - S_W(U_\mu(x)) = \mathcal{O}(X^2)$  for any Lie algebra element  $X$ , implies

$$\text{Tr} \left[ \sigma_i \left( U_\mu(x) \tilde{U}_\mu^\dagger(x) - \tilde{U}_\mu(x) U_\mu^\dagger(x) \right) \right] = 0, \quad (2.3.4)$$

where  $\sigma_i$  are the Pauli matrices. This is easily seen to imply that  $U_\mu(x) \tilde{U}_\mu^\dagger(x)$  is a multiple of the identity, and as  $\tilde{U}_\mu$  is the sum of  $SU(2)$  matrices, it can be written as  $\tilde{U}_\mu = a_0 + i\vec{a} \cdot \vec{\sigma}$ , with  $a_\mu \in \mathbb{R}^4$ . If we define  $\|\tilde{U}_\mu\| = \sqrt{a_\mu^2}$ , eq. (2.3.4) is seen to imply

$$U_\mu(x) = \pm \tilde{U}_\mu(x) / \|\tilde{U}_\mu(x)\|. \quad (2.3.5)$$

As we are only interested in stable solutions (i.e. local minima of the action), the plus sign in eq. (2.3.5) is the relevant one. The process of iteratively finding the solution to the equations of motion is called cooling [42], as in all cases it is devised such that the action is lowered after each iteration. The easiest is to simply choose  $U'_\mu(x) = \tilde{U}_\mu(x) / \|\tilde{U}_\mu(x)\|$  since a fixed point of this iteration is clearly a solution to the equations of motion. An efficient way to sweep through the lattice is to divide for each  $\mu$  the links  $U_\mu(x)$  in two mutually exclusive checkerboard patterns  $\Pi_\mu^i$  such that all links on a particular pattern  $\Pi_\mu^i$  (i.e. for fixed  $i$  and  $\mu$ ) can be changed simultaneously, which is a well-known trick to vectorize this procedure. At the cost of roughly a factor of two in memory use, vectorization is also achieved for the modified action we have considered for our numerical simulations:

$$S(\varepsilon) = \frac{4-\varepsilon}{3} \sum_{x,\mu,\nu} \text{Tr} \left( 1 - \nu \begin{array}{c} \mu \\ \downarrow \\ \mu \end{array} \right) + \frac{\varepsilon-1}{48} \sum_{x,\mu,\nu} \text{Tr} \left( 1 - \nu \begin{array}{c} \mu \\ \leftarrow \\ \mu \end{array} \right). \quad (2.3.6)$$

The meaning of the parameter  $\varepsilon$  will become clear in the next section. For ease of our numerical studies we have not considered modified single plaquette actions (see also the next section for a discussion on the adjoint and Manton actions).

## 2.4 Lattice artefacts

To calculate the effect of the discretization on the solutions of the equations of motion we first take a smooth continuum configuration (not necessarily a solution)  $A_\mu(x)$ . For definiteness we put  $L = 1$  and  $N_s$  the number of lattice points in the spatial direction, so that  $a = 1/N_s$ . We put this configuration on the lattice by defining:

$$U_\mu(x) = P \exp \left( \int_0^a ds A_\mu(x + s\hat{\mu}) \right). \quad (2.4.1)$$

The value of the plaquette thus corresponds to parallel transport around a square and can easily be proven to be given by [55] ( $\mathcal{D}_\mu = \partial_\mu + A_\mu(x)$  the covariant derivative in the fundamental representation)

$$\text{Tr} \left( \nu \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} \right) = \text{Tr} (e^{a\mathcal{D}_\mu(x)} e^{a\mathcal{D}_\nu(x)} e^{-a\mathcal{D}_\mu(x)} e^{-a\mathcal{D}_\nu(x)}). \quad (2.4.2)$$

The proof simply amounts to observing that if  $A_\mu(x) = A_\mu$ , i.e.  $A_\mu$  is space-time independent, then  $\text{Tr} \left( \nu \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} \right) = \text{Tr} (e^{aA_\mu} e^{aA_\nu} e^{-aA_\mu} e^{-aA_\nu})$  and eq. (2.4.2) is the only way to make this formula gauge invariant under arbitrary (i.e.  $x$  dependent) gauge transformations. Using the Campbell-Baker-Hausdorff formula, eq. (2.4.2) can be expressed in terms of products of covariant derivatives  $D_\mu$  (in the adjoint representation) acting on the curvature  $F_{\mu\nu} \equiv [D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ , e.g.  $D_\mu F_{\mu\nu} = [D_\mu, [D_\mu, D_\nu]]$ . As the action involves a sum over all  $x$ ,  $\mu$  and  $\nu$ , things can be considerably simplified by computing, what we will call, the clover average

$$\begin{aligned} \langle \text{Tr} \left( \nu \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} \right) \rangle_{\text{clover}} &\equiv \frac{1}{4} \text{Tr} \left( \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} \right) \\ &\equiv \frac{1}{4} \text{Tr} \left\{ e^{-a\mathcal{D}_\mu} e^{-a\mathcal{D}_\nu} e^{a\mathcal{D}_\mu} e^{a\mathcal{D}_\nu} + e^{-a\mathcal{D}_\nu} e^{a\mathcal{D}_\mu} e^{a\mathcal{D}_\nu} e^{-a\mathcal{D}_\mu} \right. \\ &\quad \left. + e^{a\mathcal{D}_\mu} e^{-a\mathcal{D}_\nu} e^{-a\mathcal{D}_\mu} e^{a\mathcal{D}_\nu} + e^{a\mathcal{D}_\nu} e^{a\mathcal{D}_\mu} e^{-a\mathcal{D}_\nu} e^{-a\mathcal{D}_\mu} \right\} \\ &= \text{Tr} \left\{ 1 + \frac{a^4}{2} F_{\mu\nu}^2(x) - \frac{a^6}{24} \left( (D_\mu F_{\mu\nu}(x))^2 + (D_\nu F_{\mu\nu}(x))^2 \right) + \frac{a^8}{24} \left[ F_{\mu\nu}^4(x) \right. \right. \\ &\quad \left. \left. + \frac{1}{30} \left( (D_\mu^2 F_{\mu\nu}(x))^2 + (D_\nu^2 F_{\mu\nu}(x))^2 \right) + \frac{1}{3} D_\mu^2 F_{\mu\nu}(x) D_\nu^2 F_{\mu\nu}(x) \right. \right. \\ &\quad \left. \left. - \frac{1}{4} (D_\mu D_\nu F_{\mu\nu}(x))^2 \right] \right\} + \mathcal{O}(a^{10}) + \text{total derivative terms}, \quad (2.4.3) \end{aligned}$$

for which the multiple Campbell-Baker-Hausdorff expansion of eq. (2.4.2) is required to  $\mathcal{O}(a^6)$ , obtained with the aid of the symbolic manipulation program FORM [56]. The clover average allows one to ignore many terms (all those odd in any of the indices) in evaluating the trace of the exponent.

Eq. (2.4.3) was also derived using the non-Abelian Stokes formula [57] ( $s_0 \equiv 1$ )

$$\begin{aligned} U_{\mu\nu}(x) &\equiv U_\mu(x) U_\nu(x + a\hat{\mu}) U_\mu^\dagger(x + a\hat{\nu}) U_\nu^\dagger(x) \\ &= P \exp \left( a^2 \int_0^1 ds \int_0^1 dt \mathcal{F}_{\mu\nu}(x + as\hat{\mu} + at\hat{\nu}) \right) \\ &\equiv 1 + \sum_{n=1}^{\infty} \prod_{i=1}^n \int_0^{s_{i-1}} ds_i \int_0^{t_{i-1}} dt_i a^2 \mathcal{F}_{\mu\nu}(x + as_i\hat{\mu} + at_i\hat{\nu}) \\ &\quad \times \dots a^2 \mathcal{F}_{\mu\nu}(x + as_n\hat{\mu} + at_n\hat{\nu}), \quad (2.4.4) \end{aligned}$$

where  $\mathcal{F}_{\mu\nu}(y)$  equals  $F_{\mu\nu}(y)$  up to the backtracking loop that connects  $y$  to  $x$ , or:

$$\begin{aligned} V(s, t) &\equiv P \exp \left( a \int_0^s d\bar{s} A_\mu(x + a\bar{s}\hat{\mu}) \right) P \exp \left( a \int_0^t d\bar{t} A_\nu(x + as\hat{\mu} + a\bar{t}\hat{\nu}) \right), \\ \mathcal{F}_{\mu\nu}(x + as\hat{\mu} + at\hat{\nu}) &\equiv V(s, t) F_{\mu\nu}(x + as\hat{\mu} + at\hat{\nu}) V^\dagger(s, t). \quad (2.4.5) \end{aligned}$$

To obtain the result of eq. (2.4.3) one now expands  $\mathcal{F}_{\mu\nu}(x + a s \hat{\mu} + a t \hat{\nu})$  around the point  $x$ , making use of the identity:

$$\partial_\mu^n \partial_\nu^m \mathcal{F}_{\mu\nu}(x) = D_\mu^n D_\nu^m F_{\mu\nu}(x). \quad (2.4.6)$$

Note that the ordering of the covariant derivatives in the right hand side of eq. (2.4.6) is essential. Also crucial is that the path ordering  $U(s, t) \equiv \text{Pexp}(\int_s^t du A(u))$  (where  $A(t) = \bar{e}_\mu A_\mu(x + t\bar{e})$  for some unit vector  $\bar{e}$ ) is compatible with the covariant derivative, i.e.  $\bar{e}_\mu \mathcal{D}_\mu(x + s\bar{e})U(s, t) = 0 = \bar{e}_\mu \mathcal{D}_\mu(x + t\bar{e})U^\dagger(s, t)$  (in this respect we have corrected the formula in ref. [57]). Inserting the Taylor expansion of  $\mathcal{F}_{\mu\nu}(x + a s \hat{\mu} + a t \hat{\nu})$  with respect to  $(s, t)$  in eq. (2.4.4), gives the result of eq. (2.4.3). A very useful check is that the symmetry implied by  $U_{\mu\nu}(x) = U_{\nu\mu}^\dagger(x)$ , not explicit at intermediate steps of the calculation, is respected by the final result.

Using eqs. (2.4.1), (2.4.3) one finds to<sup>3</sup>  $\mathcal{O}(a^8)$  for the modified action  $S(\varepsilon)$

$$\begin{aligned} S(\varepsilon) = & \sum_{x, \mu, \nu} \text{Tr} \left\{ -\frac{a^4}{2} F_{\mu\nu}^2 + \frac{\varepsilon a^6}{24} \left( (D_\mu F_{\mu\nu}(x))^2 + (D_\nu F_{\mu\nu}(x))^2 \right) \right. \\ & - \frac{(15\varepsilon - 12)a^8}{72} \left[ F_{\mu\nu}^4(x) + \frac{1}{30} \left( (D_\mu^2 F_{\mu\nu}(x))^2 + (D_\nu^2 F_{\mu\nu}(x))^2 \right) \right. \\ & \left. \left. + \frac{1}{3} D_\mu^2 F_{\mu\nu}(x) D_\nu^2 F_{\mu\nu}(x) - \frac{1}{4} (D_\mu D_\nu F_{\mu\nu}(x))^2 \right] \right\}. \quad (2.4.7) \end{aligned}$$

Obviously,  $S(\varepsilon = 1)$  corresponds to the Wilson action, and the sign of the leading lattice artefacts are simply reversed by changing the sign of  $\varepsilon$ . Most of the numerical results were obtained for  $\varepsilon = -1$ , but  $\varepsilon$  is useful in the initial cooling from a random configuration. By keeping  $\varepsilon > 0$  as long as  $S > 8\pi^2$ , and only switching to  $\varepsilon = -1$  when  $S \approx 8\pi^2$ , we can avoid the solution getting stuck at higher topological charges. Once we set  $\varepsilon = -1$ , we have yet to see an instanton fall through the lattice. We will come back to these issues when discussing the numerical results. Also note that, as  $\text{Tr}_{\text{ad}}(U) = |\text{Tr}(U)|^2 - 1$ , one finds that the Wilson action in the adjoint representation ( $S_{W, \text{ad}}$ ) satisfies  $S_{W, \text{ad}} = 4S(\varepsilon = 1) + \mathcal{O}(a^8)$  and does not allow us to change the sign of the  $a^6$  term. The same holds for the Manton action [58] which by definition agrees to  $\mathcal{O}(a^6)$  with the Wilson action.

In the past, more complicated improved actions were considered [25], for which we will present the result similar to eq. (2.4.7), as it allows us to predict whether or not they give rise to stable solutions [40]. It also allows comparison with earlier results by Lüscher and Weisz [25]. In the following, the coefficients in front of the  $c_i$  are to match with the definitions of ref. [25]. The averages  $\langle \dots \rangle$  are similar to the clover average above, implying averaging over  $\pm \hat{\mu}$ ,  $\pm \hat{\nu}$  and (for the non-planar loops)  $\pm \hat{\lambda}$ . After some algebra one finds:

<sup>3</sup>We count factors of  $a$  on the Lagrange level, i.e. before performing the summation over  $x$ .

$$\begin{aligned}
S(\{c_i\}) &\equiv \sum_x \text{Tr} \left\{ c_0 \sum_{\mu \neq \nu} \langle 1-\nu \square_{x\mu} \rangle + 2c_1 \sum_{\mu \neq \nu} \langle 1-\nu \square_{x\mu} \rangle \right. \\
&\quad \left. + \frac{4}{3}c_2 \sum_{\mu \neq \nu \neq \lambda} \langle 1-\nu \square_{x\mu\lambda} \rangle + 4c_3 \sum_{\mu \neq \nu \neq \lambda} \langle 1-\nu \square_{x\mu\lambda} \rangle \right\} \\
&= -\frac{a^4}{2}(c_0 + 8c_1 + 8c_2 + 16c_3) \sum_{x,\mu,\nu} \text{Tr} (F_{\mu\nu}(x))^2 \\
&\quad + \frac{a^6}{12}(c_0 + 20c_1 - 4c_2 + 4c_3) \sum_{x,\mu,\nu} \text{Tr} (D_\mu F_{\mu\nu}(x))^2 \\
&\quad + \frac{a^6}{3}(c_2 + 3c_3) \sum_{x,\mu,\nu,\lambda} \text{Tr} (D_\mu F_{\mu\lambda}(x) D_\nu F_{\nu\lambda}(x)) \\
&\quad + \frac{a^6}{3}c_2 \sum_{x,\mu,\nu,\lambda} \text{Tr} (D_\mu F_{\nu\lambda}(x))^2 + \mathcal{O}(a^8). \tag{2.4.8}
\end{aligned}$$

One can therefore achieve tree-level improvement by choosing [25]  $c_0 + 8c_1 + 8c_2 + 16c_3 = 1$ ,  $c_0 + 20c_1 - 4c_2 + 4c_3 = 0$  and  $c_2 = c_2 + 3c_3 = 0$ . Note that the condition  $c_2 + 3c_3 = 0$  only applies off-shell, since on-shell  $\sum_{x,\mu,\nu,\lambda} \text{Tr} (D_\mu F_{\mu\lambda}(x) D_\nu F_{\nu\lambda}(x)) = 0$ .

Iwasaki and Yoshié [40] considered cooling for the tree-level improved Lüscher-Weisz action, that is  $c_0 = \frac{2}{3}$ ,  $c_1 = -\frac{1}{12}$  and  $c_{2,3} = 0$ , for which the  $a^6$  term vanishes. The  $a^8$  term will have to be computed to settle stability. From eq. (2.4.7) one sees that the  $a^6$  term has a definite sign. This is no longer the case for the  $a^8$  term. The same holds for the Lüscher-Weisz action:

$$\begin{aligned}
S_{\text{LW}} &= \sum_{x,\mu,\nu} \text{Tr} \left[ -\frac{a^4}{2} F_{\mu\nu}^2(x) + \frac{a^8}{24} \left\{ F_{\mu\nu}^4(x) + \frac{1}{3} D_\mu^2 F_{\mu\nu}(x) D_\nu^2 F_{\mu\nu}(x) \right. \right. \\
&\quad \left. \left. - \frac{1}{4} (D_\mu D_\nu F_{\mu\nu}(x))^2 + \frac{4}{15} (D_\mu^2 F_{\mu\nu}(x))^2 \right\} \right] + \mathcal{O}(a^{10}). \tag{2.4.9}
\end{aligned}$$

To decide in these cases if the lattice admits a stable solution (i.e. its action increases with decreasing  $\rho$ ), one can compute the lattice action using explicitly the topological charge-one instanton solution with scale  $\rho$ . Eqs. (2.4.7), (2.4.8) and (2.4.9) are only valid as long as  $a \ll \rho$  because for  $\rho \approx a$  the expansion in powers of  $a$  no longer converges. For  $\rho \ll L$  to a good approximation we can substitute the infinite-volume continuum instanton solution:

$$A_\mu(x) = -i \frac{\eta_{\mu\nu}^a x_\nu \sigma_a}{(x^2 + \rho^2)}, \tag{2.4.10}$$

with  $\eta_{\mu\nu}^a$  the self-dual 't Hooft tensor [18]. When  $\rho \approx L$  the solution will of course be modified by the boundary effect. Substituting eq. (2.4.10) we find

$$\begin{aligned}
S(\epsilon) &= 8\pi^2 \left\{ 1 - \frac{\epsilon}{5} (a/\rho)^2 - \frac{15\epsilon - 12}{210} (a/\rho)^4 + \mathcal{O}((a/\rho)^6) \right\}, \\
S_{\text{LW}} &= 8\pi^2 \left\{ 1 - \frac{17}{210} (a/\rho)^4 + \mathcal{O}((a/\rho)^6) \right\}. \tag{2.4.11}
\end{aligned}$$

We thus confirm the observation of Iwasaki and Yoshié [40] that the Lüscher-Weisz action has no stable instanton solutions. Since  $S(\varepsilon = 0) = 8\pi^2 \{1 + \frac{25}{35}(a/\rho)^4 + \mathcal{O}((a/\rho)^6)\}$  we predict even at  $\varepsilon = 0$  the lattice to have stable solutions, which we have verified for the case with twisted boundary conditions in the time direction (see below).

Iwasaki and Yoshié [40] also considered cooling for Wilson's choice [22] (W) of  $c_0 = 4.376$ ,  $c_1 = -0.252$ ,  $c_2 = -0.17$  and  $c_3 = 0$  and for (R)  $c_0 = 9$ ,  $c_1 = -1$  and  $c_2 = c_3 = 0$ . To  $\mathcal{O}(a^6)$  these actions effectively correspond respectively to  $\varepsilon = -2.704$  and  $\varepsilon = -11$ , which for the case (W) we computed by substituting the continuum instanton solution. Indeed, they see stability up to 250 sweeps<sup>4</sup> in both cases.

## 2.5 Non-leading lattice artefact corrections

In presenting eq. (2.4.11) we have replaced the sum over the lattice points by an integral and ignored the fact that on the lattice the equations of motion are modified. Both effects turn out to be small, the first exponential in  $\rho/a$ , the other gives a correction to the expression for  $S(\varepsilon)$  in eq. (2.4.11) proportional to  $\varepsilon^2(a/\rho)^4$  (whereas the correction to any tree-level improved action, in particular  $S_{\text{LW}}$  and  $S(\varepsilon = 0)$ , is proportional to  $(a/\rho)^8$ ).

We wish to compute  $\sum_x f(x) = \sum_n f(na)$ , for which we can use its Fourier decomposition

$$\begin{aligned} a^4 \sum_x f(x) &= a^4 \sum_{n \in \mathbb{Z}^4} \sum_k e^{ik \cdot na} \tilde{f}(k) \\ &= a^4 N_s^3 N_t \sum_{p \in \mathbb{Z}^4} \tilde{f}\left(\frac{2\pi p}{a}\right) = \sum_{p \in \mathbb{Z}^4} \int d^4 x e^{-2\pi i p \cdot x/a} f(x). \end{aligned} \quad (2.5.1)$$

The terms with  $p \neq 0$  give the error one makes when replacing the lattice sum by an integral. For  $a \ll \rho \ll L$  and  $f(x) = -\frac{1}{2} \text{Tr}(F_{\mu\nu}^2(x + x_0))$  one finds explicitly (using eq. (2.4.10))

$$\begin{aligned} -\frac{a^4}{2} \sum_{x, \mu, \nu} \text{Tr}(F_{\mu\nu}^2(x + x_0)) &= 8\pi^2 \left[ 1 + \sum_{p \in \mathbb{Z}^4 \setminus \{0\}} 2\pi^2 p^2 (\rho/a)^2 \cos(2\pi p \cdot x_0/a) K_2(2\pi |p| \rho/a) \right] \\ &= 8\pi^2 \left[ 1 - 8\pi^2 (\rho/a)^{3/2} e^{-2\pi \rho/a} (1 + \mathcal{O}(a/\rho)) \right] \end{aligned} \quad (2.5.2)$$

(with  $K_2$  the modified Bessel function [59]). Here we have taken  $x_0$  to coincide with a point on the dual lattice,  $2x_0^\mu = a$  for all  $\mu$ , as this minimizes the action.

To estimate the shift in the equations of motion due to the lattice artefacts we again consider  $\rho \gg a$ , so that in a good approximation the action is given by

<sup>4</sup>A sweep is an iteration step in which all links are updated once.

$$\begin{aligned} \bar{S}(\varepsilon, A_\mu) = & \sum_{\mu,\nu} \int d^4x \left\{ -\frac{1}{2} \text{Tr} (F_{\mu\nu}(x))^2 + \frac{\varepsilon a^2}{12} \text{Tr} (D_\mu F_{\mu\nu}(x))^2 \right. \\ & - a^4 \frac{(15\varepsilon - 12)}{72} \text{Tr} \left[ (F_{\mu\nu}(x))^4 + \frac{1}{15} (D_\mu^2 F_{\mu\nu}(x))^2 \right. \\ & \left. \left. + \frac{1}{3} (D_\mu^2 F_{\mu\nu}(x) D_\nu^2 F_{\mu\nu}(x)) - \frac{1}{4} (D_\mu D_\nu F_{\mu\nu}(x))^2 \right] \right\} + \mathcal{O}(a^6), \end{aligned} \quad (2.5.3)$$

which implies the equations of motion:

$$\begin{aligned} \sum_\nu D_\nu F_{\nu\mu} &= \varepsilon a^2 H_\mu + \mathcal{O}(a^4), \\ H_\mu &\equiv -\sum_\nu \left( \frac{1}{12} D_\nu^3 F_{\nu\mu} + \frac{1}{6} [F_{\mu\nu}, D_\mu F_{\mu\nu}] + \frac{1}{12} D_\mu^2 D_\nu F_{\nu\mu} \right). \end{aligned} \quad (2.5.4)$$

As eq. (2.5.3) breaks the scale invariance, there will in general not be solutions close to eq. (2.4.10). Variation with respect to  $\rho$  no longer leaves the action invariant. Still, since this variation corresponds to a near zero-mode, it makes sense to expect quasi-stability under cooling. The action changes only slowly in the direction of this near zero-mode but is predominantly lowered in those directions that leave the curvature square integrable and are spanned by the non-zero modes of the quadratic fluctuation operator for the action, which in the background gauge corresponds to

$$\mathcal{M}_{\lambda\sigma} = \delta_{\lambda\sigma} D_\mu^2 + 2adF_{\lambda\sigma}. \quad (2.5.5)$$

If  $\mathcal{P}$  is the projection operator on the normalizable non-zero modes of  $\mathcal{M}$  one has (at  $\rho = 1$ )

$$\begin{aligned} A_\mu^{(\varepsilon)} &= A_\mu^{(0)} + \varepsilon a^2 \mathcal{P} \mathcal{M}_{\mu\nu}^{-1} \mathcal{P} H_\nu + \mathcal{O}(a^4) \equiv A_\mu^{(0)} + \varepsilon a^2 A_\mu^{(1)} + \mathcal{O}(a^4), \\ H_\mu(x) &= \frac{16}{(1+x^2)^5} \psi_\mu^{(\frac{3}{2}, \frac{3}{2})}(x) + \frac{8}{(1+x^2)^5} \psi_\mu^{(\frac{1}{2}, \frac{1}{2})}(x), \end{aligned} \quad (2.5.6)$$

where  $H_\mu(x)$  is evaluated by substituting for  $A_\mu^{(0)}(x)$  the continuum solution  $A_\mu(x)$  given in eq. (2.4.10). For convenience we introduced the quantities  $\psi_\mu^{(i,j)}(x)$ :

$$\begin{aligned} \psi_\mu^{(\frac{1}{2}, \frac{1}{2})}(x) &= i \sum_{\nu, \alpha} \eta_{\mu\nu}^\alpha x_\nu \sigma_\alpha, \quad \psi_\mu^{(\frac{3}{2}, \frac{3}{2})}(x) = i \sum_{\nu, \alpha} \eta_{\mu\nu}^\alpha x_\nu \sigma_\alpha (x^2 - 6x_\mu^2), \\ \psi_\mu^{(\frac{3}{2}, \frac{1}{2})}(x) &= i \sum_{\nu, \alpha} \eta_{\mu\nu}^\alpha x_\nu \sigma_\alpha (3x^2 - 3x_\mu^2 - 5x_\nu^2), \end{aligned} \quad (2.5.7)$$

which are eigenfunctions of the angular momentum operators  $\bar{L}_i^2$  and  $\bar{J}^2$  (as defined in [18]  $L_i^\alpha = -\frac{i}{2} \eta_{\mu\nu}^\alpha x_\mu \partial_\nu$ ,  $J^\alpha = L_i^\alpha + \text{ad}(\frac{\sigma_\alpha}{2})$ ). To compute  $\mathcal{M}_{\mu\nu}^{-1} \mathcal{P} H_\nu$  one can use for  $\mathcal{M}_{\mu\nu}^{-1}$  the explicit expression [60]  $\mathcal{M}_{\mu\nu}^{-1} \equiv \bar{\eta}_{\mu\lambda}^\alpha D_\lambda (D_\alpha^{-2}) \bar{\eta}_{\nu\sigma}^\alpha D_\sigma$  (in the gauge  $D_\mu(\mathcal{P} H_\mu) = 0$ ). The result, which can be verified by applying  $\mathcal{M}_{\mu\nu}$ , is found to be

$$\begin{aligned}
\mathcal{M}_{\mu\nu}^{-1}\mathcal{P}H_\nu &= \left( \frac{\ln(1+x^2)}{5(1+x^2)^2} - \frac{1}{3(1+x^2)^3} + \frac{1}{10(1+x^2)} \right) \psi_{\mu}^{(\frac{1}{2}, \frac{1}{2})}(x) \\
&+ \left( \frac{2\ln(1+x^2)}{5x^8(1+x^2)^2} - \frac{6+3x^2-x^4}{15x^6(1+x^2)^3} \right) \psi_{\mu}^{(\frac{3}{2}, \frac{3}{2})}(x) \\
&+ \left( \frac{2(3+5x^2)\ln(1+x^2)}{5x^8(1+x^2)^2} - \frac{6+13x^2+4x^4}{5x^6(1+x^2)^3} \right) \psi_{\mu}^{(\frac{5}{2}, \frac{5}{2})}(x). \quad (2.5.8)
\end{aligned}$$

The algebraic manipulation program Mathematica [61] was useful in obtaining and checking these results. Despite its appearance, this result is regular at  $x \rightarrow 0$ . However, it contains a non-normalizable deformation (since  $\psi_{\mu}^{(1/2, 1/2)}(x)/(1+x^2) = -A_{\mu}^{(0)}(x)$ ), which would make the action diverge and should be removed by projecting on normalizable deformations:

$$\begin{aligned}
A_{\mu}^{(1)} &= \mathcal{P}\mathcal{M}_{\mu\nu}^{-1}\mathcal{P}H_\nu = \mathcal{M}_{\mu\nu}^{-1}\mathcal{P}H_\nu - \frac{(23+17x^2)}{60(1+x^2)^2} \psi_{\mu}^{(\frac{1}{2}, \frac{1}{2})}(x), \\
\mathcal{M}_{\mu\nu}A_{\nu}^{(1)} &= H_{\mu} - \frac{8\psi_{\mu}^{(\frac{1}{2}, \frac{1}{2})}(x)}{5(1+x^2)^3}. \quad (2.5.9)
\end{aligned}$$

One easily verifies that  $A_{\mu}^{(1)}$  and  $\mathcal{M}_{\mu\nu}A_{\nu}^{(1)}$  are both square integrable and orthogonal to the zero-mode  $\partial A_{\mu}^{(0)}/\partial\rho|_{\rho=1} = 2\psi_{\mu}^{(1/2, 1/2)}(x)/(1+x^2)^2$ .

We can now substitute  $A_{\mu}^{(\varepsilon)} = A_{\mu}^{(0)} + \varepsilon a^2 A_{\mu}^{(1)} + \mathcal{O}(a^4)$  in eq. (2.4.7) to obtain the shift in the action. It will be useful for verifying eq. (2.6.2) if we evaluate

$$\begin{aligned}
\bar{S}(\bar{\varepsilon}, A_{\mu}^{(\varepsilon)}) &= \bar{S}(\bar{\varepsilon}, A_{\mu}^{(0)}) - 2\varepsilon\bar{\varepsilon}a^4 \int d^4x \text{Tr}(A_{\mu}^{(1)}(x)H_{\mu}(x)) \\
&+ \varepsilon^2 a^4 \int d^4x \text{Tr}(A_{\mu}^{(1)}(x)\mathcal{M}_{\mu\nu}A_{\nu}^{(1)}(x)) + \mathcal{O}(a^6). \quad (2.5.10)
\end{aligned}$$

The equations of motion for  $A_{\mu}^{(0)}$  were used to simplify the term linear in the shift  $A_{\mu}^{(\varepsilon)} - A_{\mu}^{(0)} = \varepsilon a^2 A_{\mu}^{(1)} + \mathcal{O}(a^4)$ , which makes it evident that the  $\mathcal{O}(a^4)$  term in  $A_{\mu}^{(\varepsilon)}$  will only contribute  $\mathcal{O}(a^6)$  to eq. (2.5.10). To evaluate the action used for cooling one simply equates  $\bar{\varepsilon}$  to  $\varepsilon$ . Evaluating the integrals, reintroducing the  $\rho$  dependence using trivial dimensional arguments, gives

$$\bar{S}(\bar{\varepsilon}, A_{\mu}^{(\varepsilon)}) = 8\pi^2 \left\{ 1 - \frac{\bar{\varepsilon}}{5} \left( \frac{a}{\rho} \right)^2 - \left[ \frac{15\bar{\varepsilon} - 12}{210} + \frac{284\varepsilon\bar{\varepsilon}}{2625} - \frac{179\varepsilon^2}{5250} \right] \left( \frac{a}{\rho} \right)^4 + \mathcal{O}\left( \left( \frac{a}{\rho} \right)^6 \right) \right\}. \quad (2.5.11)$$

At  $\varepsilon = \bar{\varepsilon} = -1$  the shift due to the modified equations of motion in the  $(a/\rho)^4$  term is 58%.

Strictly speaking our expression for the  $\rho$  dependence of the lattice action is only valid for  $\rho/a \gtrsim 2$  and  $\rho \ll L$ , since we are using the continuum infinite volume solution as the zero-order approximation. But even for  $\rho \approx L/2$  it is not unreasonable to expect the order of magnitude of the corrections to be given by eqs. (2.5.2) and (2.5.11).

## 2.6 Numerical results and discussion

This section discusses the numerical results obtained by the method of section 2.3, mainly to illustrate the viability of our ideas. For a more extensive and careful analysis see ref. [62]. We will also draw some conclusions concerning the instanton structure on  $T^3 \times \mathbb{R}$ .

We used lattices of size  $N_s^3 \times N_t$ , with  $N_s = 7$  or  $8$  and with  $N_t = 3N_s$ . At  $\varepsilon = -1$  (see eq. (2.3.6)) cooling settles to an action near  $8\pi^2$ , and we have seen stability for up to 6000 sweeps. The same is true for  $\varepsilon = 0$  in the presence of a twist (without twist our configuration ultimately decays to the vacuum at  $\varepsilon = 0$ , but remember that in that case there are no regular instanton solutions). Apart from the total action we computed separately the sum over the  $n \times m$  plaquettes, denoted by  $S_{n \times m}$  (also averaging over the two orientations if  $n \neq m$ ).  $S_{n \times m}$  is normalized by dividing by  $8\pi^2 n^2 m^2$ , so that for an infinite lattice and  $\rho/a \rightarrow \infty$ ,  $S_{n \times m} \rightarrow 1$ .

When we perform cooling with the action of eq. (2.3.6) we should take  $A_\mu = A_\mu^{(0)} + \varepsilon a^2 A_\mu^{(1)}$  in order to obtain an analytic prediction for  $S_{n \times m}$ . Eq. (2.4.3) for  $S_{1 \times 1}$  easily leads to the general result for  $S_{n \times m}$  by inserting for each index  $\mu(\nu)$  a factor  $n(m)$ . Together with eq. (2.5.10) one deduces that, to  $\mathcal{O}(a^4)$ ,

$$S_{n \times m} = 1 - \frac{(n^2 + m^2)}{2} \alpha a^2 - \left( m^2 n^2 \beta_1 - \frac{m^4 + n^4}{2} \beta_2 + \frac{n^2 + m^2}{2} \varepsilon \gamma - \varepsilon^2 \delta \right) a^4, \quad (2.6.1)$$

up to the discretization error implied by eq. (2.5.2), which for the lattices we are considering can be estimated to be not bigger than  $10^{-6}$ . This formula holds for sufficiently smooth configurations, i.e.  $\alpha a^2 \ll 1$ , even if the configuration has non-vanishing action over the entire spatial volume. It is these configurations that are of interest to us and which deviate considerably from localized instantons (eq. (2.4.10)) for which  $\rho \ll L$ . From eqs. (2.4.11), (2.5.11) we easily deduce for those localized instantons the results

$$\begin{aligned} \alpha a^2 &= \frac{1}{5} \left(\frac{a}{\rho}\right)^2, & \beta_1 a^4 &= \frac{29}{630} \left(\frac{a}{\rho}\right)^4, & \beta_2 a^4 &= \frac{2}{63} \left(\frac{a}{\rho}\right)^4, \\ \gamma a^4 &= \frac{284}{2625} \left(\frac{a}{\rho}\right)^4, & \delta a^4 &= \frac{179}{5250} \left(\frac{a}{\rho}\right)^4. \end{aligned} \quad (2.6.2)$$

From the numerical results we have obtained  $S_{1 \times 1}$ ,  $S_{1 \times 2}$ ,  $S_{2 \times 2}$  and  $S_{1 \times 3}$  on two lattices of size respectively  $7^3 \times 21$  and  $8^3 \times 24$ , for  $\varepsilon = 0$  and  $\varepsilon = -1$ , and with a twist  $n_{0i} = (1, 1, 1)$  (see table 2-1). From these we extract the coefficients in eq. (2.6.1), whose values are summarized in table 2-2 (the error due to neglecting the  $\mathcal{O}(a^6)$  term is of the order of  $(n^6 + m^6)(\alpha a^2)^3$ ).

It is interesting to analyze the untwisted case in more detail to illustrate the difficulty in having self-dual solutions at finite  $T$ . In fig. 2-1 we plot (a) the total electric and magnetic energies  $\mathcal{E}_{E,B}(t)$ , (b) the Polyakov-line  $P_1(t)$  through two particular points  $\mathbf{x}$ , and similarly for  $P_{2,3}(t)$  in (c,d). We see two features that are intimately related. First, where  $\mathcal{E}_B(t) \rightarrow 0$  the electric energy  $\mathcal{E}_E(t) \rightarrow \text{const}$ . Second, for

$N_s \times N_t$	$\epsilon$	$S_{1 \times 1}$	$S_{1 \times 2}$	$S_{2 \times 2}$	$S_{1 \times 3}$
$7^3 \times 21$	-1	0.982591	0.957050	0.928823	0.918105
$7^3 \times 21$	0	0.982287	0.956437	0.927908	0.917109
$8^3 \times 24$	-1	0.986720	0.967122	0.945887	0.936619
$8^3 \times 24$	0	0.986529	0.966736	0.945310	0.935976

Table 2-1. Numerical results obtained by cooling with  $S(\epsilon)$  and twist  $n_{0i} = (1, 1, 1)$ .

$N_s \times N_t$	$\alpha a^2$	$\beta_1/\alpha^2$	$\beta_2/\alpha^2$	$\gamma/\alpha^2$	$\delta/\alpha^2$
$7^3 \times 21$	0.01761	0.96	0.63	0.66	0.32
$8^3 \times 24$	0.01340	1.01	0.64	0.71	0.35
$\rho \ll L$	$0.2(a/\rho)^2$	1.15	0.79	2.70	0.85

Table 2-2. Coefficients appearing in eq. (2.6.1) extracted from the numerical results in table 2-1, using  $S_{1 \times 1}$ ,  $S_{2 \times 2}$  and  $S_{1 \times 2}$  (the latter at  $\epsilon = -1$  only).

the same  $t$  values where this occurs  $C_i(t)$  ( $P_i(t) = \cos(C_i(t)/2)$ ) is linear in  $t$  and independent of  $x$ . These are precisely the equations of motion when restricting to the vacuum valley. Classical motion in this valley, which itself has the geometry of a three-torus, is free. On the lattice this motion is described by the action

$$\sum_t 4N_s^3 \left\{ 1 - \cos \left( \frac{C_k(t+1) - C_k(t)}{2N_s} \right) \right\}. \quad (2.6.3)$$

One easily checks that the values of  $C_k(t+1) - C_k(t)$  obtained from figs. 2-1(b-d) quite accurately reproduce through eq. (2.6.3) the value for  $\mathcal{E}_E(t)$ . Clearly the electric tail destroys the self-duality.

The tail can be explained by the following argument. Suppose that at  $T \rightarrow \infty$  the solution describes tunneling from  $C_i^-$  to  $C_i^+$  and  $C_i^+ \neq C_i^-$ , then at finite  $T$  the periodic boundary conditions force  $C_i(t)$  to linearly interpolate between  $C_i^+$  and  $C_i^-$  over a time  $T - T_0$ , if  $T_0$  is the time interval over which  $\mathcal{E}_E(t) \neq 0$ . Thus the action, even in the continuum, would be bigger than  $8\pi^2$  by a number proportional to  $1/(T - T_0)$  except when there are solutions with  $C_i^- = C_i^+$  for  $T \rightarrow \infty$ . These can certainly not be excluded, in particular as  $C_i^+ = C_i^- = \pi$  is compatible with a twist  $n_{0i} = 1$ , but if these are very localized instantons, the lattice artefacts might make its action so big that the lattice will prefer the least localized solutions with  $C_i^+ \neq C_i^-$ . If  $T$  is not big enough the lattice will find a compromise between these two cases.

Now let us discuss the case with twist  $n_{0i} = (1, 1, 1)$ . We compare for  $N_s = 7$  and 8 (after appropriate scaling with  $N_s$ ) in fig. 2-2(a) the electric and magnetic energy profiles, and in figs. 2-2(b-d) the values of  $C_i \equiv 2\arccos(P_i)$  at the spatial lattice point with maximal energy (to be precise, with maximal  $E_i^2$ ). From this we deduce

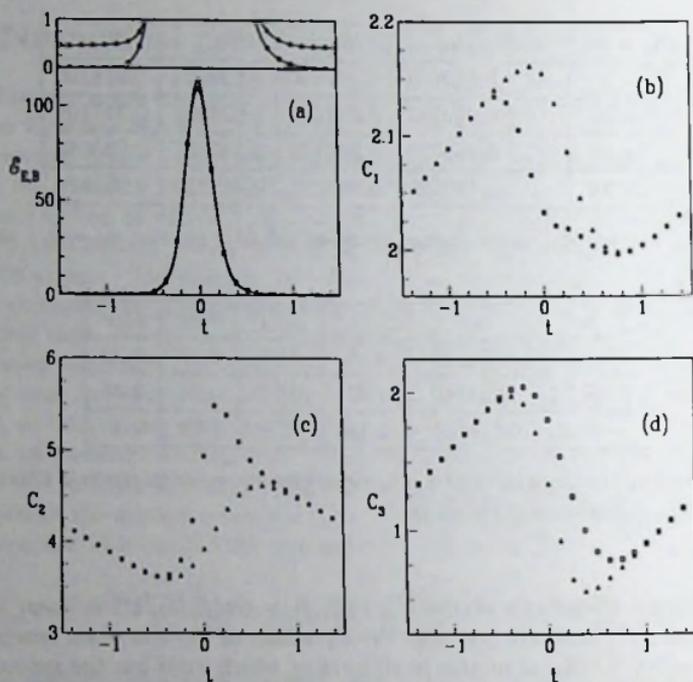


Figure 2-1. Numerical results (after scaling appropriately with  $N_s$ ) for the case of an  $8^3 \times 24$  lattice without twist, obtained from over-improved cooling at  $\epsilon = -1$ . In (a) the electric ( $\mathcal{E}_E(t)$  triangles) and magnetic ( $\mathcal{E}_B(t)$  squares) energies are plotted. In the upper part of this figure the tails are plotted at an enlarged scale. In (b-d) are plotted  $C_i(t) \equiv 2\arccos(P_i(t))$  through two distinct spatial points on the lattice.

$\mathcal{E}_E = \mathcal{E}_B$  to a high accuracy, consistent with self-duality, and an excellent scaling with  $N_s$ . The results in fig. 2-2 are obtained after roughly 6600 cooling sweeps (at  $\epsilon = -1$ ), which is necessary since the dependence of the lattice action on  $C_i^\pm$  is rather weak (at  $\epsilon = 0$  too weak to observe) and the configuration only slowly reaches the minimum of the lattice action. We have verified that the approach to this minimum is exponential, as is illustrated in fig. 2-3 where we plot the total action and the maximum of  $\mathcal{E}_B(t)$  (i.e.  $\mathcal{E}_B(0)$ ) as a function of the number of cooling sweeps. We see indeed that the maximal energy along the tunneling path decreases under cooling, which is mainly due to the increasing size, as otherwise the action should depend more strongly on the number of cooling sweeps. (For the Wilson action one sees a dramatic increase of  $\mathcal{E}_B(0)$  under cooling, until the action suddenly drops to zero.) With boundary conditions that fix the link variables at  $t = 0$  and  $t = T$  to the vacuum configurations, the approach to the minimum action is much faster.

Fixed boundary conditions were used extensively in ref. [62]. For the twisted case

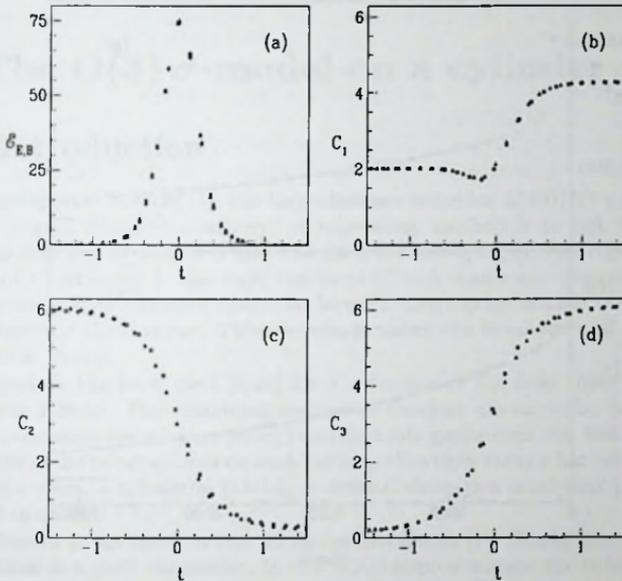


Figure 2-2. Numerical results (after scaling appropriately with  $N_s$ ) for the cases of a  $7^3 \times 21$  (squares) and an  $8^3 \times 24$  (triangles) lattice with twist  $n_{0i} = (1, 1, 1)$ , obtained from over-improved cooling at  $\epsilon = -1$ . Figure (a) contains four data sets. Two for  $\mathcal{E}_B(t)$  with the above mentioned symbols and two (crosses for  $N_s = 7$  and stars for  $N_s = 8$ ) for  $\mathcal{E}_B(t)$ . Figures (b-d) exhibit  $C_i(t)$ , through the spatial lattice point with maximal  $E_i^2$  at  $t = 0$

$n_{0i} = (1, 1, 1)$  (again for large  $T$ :  $N_t = 3N_s$  and  $N_t = 6N_s$ ) it was shown numerically that the largest instanton satisfies very special boundary conditions, namely<sup>5</sup> (for  $T \rightarrow \infty$ )  $P_i(t = T) = -P_i(t = 0) = \pm 1$ . This can also be seen from figs. 2-2(c,d). However, fig. 2-2(b) shows a deviation. The reason is that the instanton for  $P_i(t = T) = -P_i(t = 0) = \pm 1$  suffers from extra finite- $T$  effects that increase the action [62]. In ref. [62] it was also established that the largest instanton goes through a configuration with exactly one unstable mode, which thus is likely to be a sphaleron as conjectured in section 2.1. (One cannot conclude with absolute certainty that it is a sphaleron, because the precise definition [13, 14] of a sphaleron requires a global mini-max procedure that is numerically inaccessible, or at least very expensive).

One question remains open. From our results one cannot conclude whether or

<sup>5</sup>These are precisely the points in the vacuum valley around which the potential is quartic instead of quadratic, i.e. where the quantum mechanical wave functions of low-lying states will be peaked (at least for small volumes  $[0, L]^3$ ) [8].

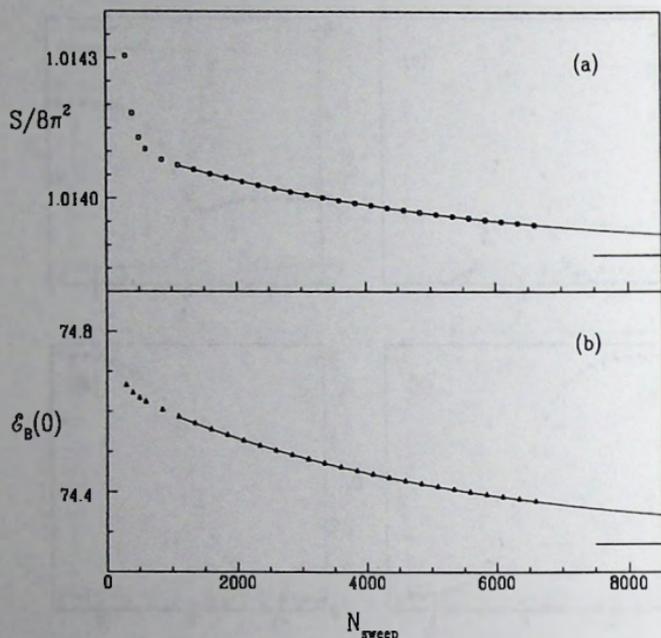


Figure 2-3. The history of the action  $S(\epsilon = -1)$  and the maximal magnetic energy  $\mathcal{E}_B(t = 0)$  as a function of the number of cooling sweeps for an  $8^3 \times 24$  lattice with twist  $\pi_{0i} = (1, 1, 1)$ , together with their exponential fits. The short lines on the right indicate the asymptotic values following from these fits.

not instanton solutions exist for  $T \rightarrow \infty$  when the twisted boundary conditions are relaxed. For the untwisted case our analysis shows that near-instantons, with action infinitesimally close to  $8\pi^2$ , exist in this limit, but this observation alone is not sufficient to conclude that these are exact solutions to the equations of motion. We think it is worthwhile to analyze the situation in the 2-dimensional  $O(3)$   $\sigma$ -model, which shares some important features with  $SU(N)$  pure gauge theory. This will be the topic of the next chapter.

As a final remark on over-improvement, let us mention that its use need not be limited to classical solutions. It may be used [63] as an efficient tool to extract topological signals from Monte Carlo data, as the action to generate a statistical ensemble need not be the same as the one used to remove the noise.

## 3 The $O(3)$ $\sigma$ -model on a cylinder

### 3.1 Introduction

Due to asymptotic freedom [1], the large-distance behavior of  $SU(N)$  gauge theories must be treated non-perturbatively. A convenient method is to put the model in a finite spatial box of length  $L$  and calculate the low-lying energy eigenstates as a function of  $L$ . At small  $L$ , the wave functions of such states are concentrated in the vacua of *classical* configuration space. At larger  $L$  they can spread out over low energy barriers between these vacua. This spreading causes the breakdown of conventional perturbation theory.

This picture has been used [8, 16] for a reduction of the field theory to a finite-dimensional system. The remaining degrees of freedom are expected to correspond to the set of vacua (or vacuum valley) and suitable paths over the barriers between vacua. One of the requirements on such paths is that they cross a barrier at its lowest point. This point, a sphaleron [13, 14], is defined through a mini-max procedure, as described in section 1.1.

Well-known paths between two vacua are instantons [17, 19, 64], interpreting (Euclidean) time as a path parameter. In the WKB approximation the tunneling amplitude is dominated by paths near instantons. Therefore it seems natural to assume that amongst the instanton paths there is one which crosses a barrier at its lowest point, i.e. it goes through a sphaleron. Let us assume that the instanton moduli-space has a scale parameter  $\rho$ . Since all instantons have the same action, a dimensional argument easily shows that the maximal energy along the tunneling path goes as  $1/\rho$ . Hence it is tempting to conjecture that instantons with maximal width, as set by  $L$ , go through a sphaleron. This was one of the motivations for the research described in the previous chapter.

The conjecture is *in general* not true. For example, consider the two-dimensional potential  $V(q_1, q_2) = (q_1^2 - 1)^2[(q_2^2 - 1)^2(2 + q_2^2) + 1] + (1 + b^2/q_2^2)^{-1}$  for small values of the parameter  $b$ . It has vacua at  $(\pm 1, 0)$  with zero energy and (to leading order in  $b$ ) sphalerons at  $(0, \pm 1)$  with energy 2. However, for small enough  $b$  the instanton does not go through a sphaleron. Instead, it goes straight through the saddle point at  $(0, 0)$ , which has one unstable mode and energy 3. Quantum mechanically, this leads to the following effect. For energy eigenvalues  $E \ll 2$ , the spreading of the wave function between the vacua is in the  $(0, 0)$  direction. However, for  $E \approx 2$  this tunneling effect is overwhelmed by a spreading over the sphalerons due to classically allowed motion. Only for  $E \gtrsim 3$  the instanton direction benefits from a similar effect.

The above clearly shows that in field theory the conjecture needs to be checked. In the previous chapter a numerical procedure was used to find the widest instantons for  $SU(2)$  gauge theory on a space-time  $T^3 \times [0, T]$  (at  $T \rightarrow \infty$ ). Numerical investi-

gations [62] have shown that the widest instanton goes through a saddle point with one unstable mode, like for  $SU(2)$  on a space-time  $S^3 \times \mathbb{R}$  [16]. It is likely these saddle points are sphalerons. In this chapter we study the two-dimensional  $O(3)$   $\sigma$ -model. We will construct (for  $T = \infty$ ) all instanton solutions, and prove rigorously that the widest instantons go through sphalerons.

It is well known that the  $O(3)$   $\sigma$ -model shares with the four-dimensional gauge theories features like renormalizability and asymptotic freedom [65]. Another similarity is still more important to us now. Both models, when put in a spatial cube with periodic boundary conditions (i.e.  $T^3$  resp.  $T^1 \equiv S^1$ ), have vacuum valleys of non-zero dimension. This can increase the number of instantons, as they must have endpoints in the vacuum valley. Like for gauge theories on  $T^3 \times T^1$  (see the previous chapter), also the  $O(3)$   $\sigma$ -model on  $T^1 \times T^1$  has no exact charge 1 instanton solutions [44, 66, 67]. This does not rule out the existence of instantons on an infinite time interval. It merely suggests the vacuum valley cannot be reached in a finite amount of time.

In order to avoid confusion we stress the following. We argued before that the maximal energy along a tunneling path decreases as  $1/\rho$  with increasing scale parameter  $\rho$ . On  $\mathbb{R}^2$  this scale parameter can be arbitrarily large, so that no sphalerons exist. A finite (spatial) volume is essential for the existence of a sphaleron. In the context of high-temperature baryon number violation, explicit scale symmetry breaking terms were added to the  $O(3)$   $\sigma$ -model [68] (in analogy to the coupling of the gauge field to the Higgs field in the electroweak sector [13]). This allows one to find sphalerons in an infinite volume, but these are completely different from, and address different physics as, the sphalerons stabilized by the finite volume cutoff.

The outline of this chapter is as follows. In section 3.2 we set up a convenient formalism, which is used in section 3.3 to derive all static solutions of the Euler-Lagrange equations. In particular we find the vacuum valley and the sphaleron solutions. Section 3.4 is devoted to constructing all instanton solutions and interpreting their moduli space. A scale parameter will emerge that relates the instanton field at  $t = +\infty$  to that at  $t = -\infty$ . We conclude in section 3.5 by putting our results in perspective with respect to the four-dimensional  $SU(2)$  case.

## 3.2 The $O(3)$ $\sigma$ -model in general coordinates

The action for the  $O(3)$   $\sigma$ -model on a cylinder reads

$$S[\vec{n}] = \frac{1}{2} \int d^2x |\partial_\mu \vec{n}(x)|^2, \quad \vec{n}(x) \in \mathbb{R}^3, \quad |\vec{n}(x)|^2 = 1, \quad (3.2.1)$$

where the integration runs over space-time  $\{(x_1, x_2) \in \mathbb{R} \times T^1\}$  ( $x_1$  being time). We use overall scale invariance to fix the length of the spatial 1-torus  $T^1$  (the circle) to be  $2\pi$ , so that  $\vec{n}(x + 2\pi\hat{e}_2) = \vec{n}(x)$ . The metric on space-time is Euclidean, and we use the summation convention over repeated indices throughout.

By definition,  $\vec{n}(x) \in S^2$ . If we use coordinates  $v^i$  ( $i = 1, 2$ ) and a metric  $g_{ij}$  on  $S^2$ , eq. (3.2.1) can be rewritten as

$$S[v] = \frac{1}{2} \int d^2x g_{ij}(v(x)) \partial_\mu v^i(x) \partial_\nu v^j(x). \quad (3.2.2)$$

In this chapter we will use both spherical coordinates  $(\vartheta, \varphi)$  and stereographic projection  $(u_1, u_2)$  which will be paired as  $u \equiv u_1 + iu_2$ , with complex conjugate  $\bar{u} = u_1 - iu_2$  (cf. ref. [64, 66]):

$$\vec{n} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = \frac{1}{1 + |u|^2} \begin{pmatrix} u + \bar{u} \\ (u - \bar{u})/i \\ |u|^2 - 1 \end{pmatrix} \quad (3.2.3)$$

(hence  $u = \cot \frac{1}{2} \vartheta e^{i\varphi}$ ).

In section 3.3 we need the Euler-Lagrange equations and the Hessian of eq. (3.2.2). A straightforward computation [69] gives (for vanishing  $\delta v(x_1 \rightarrow \pm\infty, x_2)$ ):

$$\begin{aligned} S[v + \delta v] &= S[v] + S^{(1)}[v, \xi] + S^{(2)}[v, \xi] + \mathcal{O}(\xi^3), \\ S^{(1)}[v, \xi] &= - \int d^2x (D_\mu \partial_\nu v)_i \xi^i, \\ S^{(2)}[v, \xi] &= \int d^2x \xi^i \mathcal{H}_{ij}[v] \xi^j, \\ \mathcal{H}_{ij}[v] &= -\frac{1}{2} (D_\mu D_\mu)_{ij} + \frac{1}{2} R_{kijl} \partial_\mu v^k \partial_\nu v^l. \end{aligned} \quad (3.2.4)$$

Here  $\xi = \delta v + \mathcal{O}(\delta v^2)$  is defined in such a way that it transforms covariantly, see ref. [69] for details. Since the action is a scalar, this guarantees that  $\mathcal{H}$  is a tensor, as can be verified from the explicit form. Furthermore,  $D_\mu$  is the covariant derivative and  $R$  is the Riemann tensor, both at the point  $v$ :

$$\begin{aligned} (D_\mu \lambda)_i &= \partial_\mu \lambda_i - \Gamma^j_{ik} \partial_\mu v^k \lambda_j, \quad (\text{for any vector } \lambda), \\ \Gamma_{kij} &= \frac{1}{2} (\hat{\partial}_j g_{ik} + \hat{\partial}_i g_{jk} - \hat{\partial}_k g_{ij}), \quad (\hat{\partial}_i \equiv \frac{\partial}{\partial v^i}), \\ R^l_{ijk} &= \hat{\partial}_j \Gamma^l_{ik} - \hat{\partial}_k \Gamma^l_{ij} + \Gamma^m_{ik} \Gamma^l_{mj} - \Gamma^m_{ij} \Gamma^l_{mk}. \end{aligned} \quad (3.2.5)$$

The Euclidean action, eq. (3.2.2), equals precisely half the total area that  $v(T^1 \times \mathbb{R})$  covers on  $S^2$ . This action also naturally occurs in string theory (see e.g. ref. [70]). Therefore, extremizing the action (which amounts to putting  $D_\mu \partial_\nu v = 0$ , eq. (3.2.4)) gives a geodesic surface (in affine parametrization) on  $S^2$ .

The winding number  $Q[v]$  measures the number of times  $T^1 \times \mathbb{R}$  is wrapped around  $S^2$  by  $v$  (therefore it is invariant under continuous deformations of  $v$ ). With  $\varepsilon_{12} = -\varepsilon_{21} = 1$ ,  $\varepsilon_{11} = \varepsilon_{22} = 0$  its definition reads

$$\begin{aligned} Q[v] &= -\frac{1}{8\pi} \int d^2x \varepsilon_{\mu\nu} \vec{n} \cdot (\partial_\mu \vec{n} \times \partial_\nu \vec{n}) \\ &= \frac{1}{8\pi} \int d^2x \varepsilon_{\mu\nu} \partial_\mu \vec{n} \cdot (\vec{n} \times \partial_\nu \vec{n}) = \frac{1}{8\pi} \int d^2x \varepsilon_{\mu\nu} g_{ij} \partial_\mu v^i \partial_\nu v^j, \end{aligned} \quad (3.2.6)$$

where  $\lambda_\nu^j$  is the vector in the tangent space of  $S^2$  (at the point  $v$ ) corresponding to  $\bar{n} \times \partial_\nu \bar{n}$ ; up to orientation  $\lambda_\nu$  is defined by  $g_{ij} \lambda_\nu^i \lambda_\nu^j = g_{ij} \partial_\nu v^i \partial_\nu v^j$ ,  $g_{ij} \lambda_\nu^i \partial_\nu v^j = 0$  (no summations over  $\nu$ ). From this follows the well-known formula [64]

$$S = \bar{S}_\pm \mp 4\pi Q, \quad \bar{S}_\pm[v] = \frac{1}{4} \int d^2x g_{ij} (\partial_\mu v^i \pm \varepsilon_{\mu\nu} \lambda_\nu^i) (\partial_\mu v^j \pm \varepsilon_{\mu\rho} \lambda_\rho^j). \quad (3.2.7)$$

We define an instanton to be a minimal-action configuration in the sector  $Q = +1$ . Therefore it has action  $4\pi$  and satisfies the instanton equation

$$\partial_\mu v^i = \varepsilon_{\mu\nu} \lambda_\nu^i. \quad (3.2.8)$$

### 3.3 Sphalerons, vacua and other static solutions

In this section we study the potential, or energy functional, which is the action (3.2.2) restricted to space, i.e. without time-dependence and without integration over time:

$$V[v] = \frac{1}{2} \int_{T^1} dx_2 g_{ij}(v) \partial_2 v^i \partial_2 v^j, \quad (v = v(x_2)). \quad (3.3.1)$$

This is the geodesic action for a curve  $v(x_2)$  on  $S^2$ . Therefore all static solutions of the Euler-Lagrange equations, being extrema of eq. (3.3.1), are big circles on  $S^2$  (affinely parametrized by  $x_2$ ). Remembering the requirement  $v(x_2 + 2\pi) = v(x_2)$  we conclude that the most general static solution reads (after an SO(3) rotation)

$$\bar{n}_k(x_2) = \begin{pmatrix} \cos(kx_2) \\ \sin(kx_2) \\ 0 \end{pmatrix}, \quad k \in \mathbb{Z}. \quad (3.3.2)$$

One easily computes the energy

$$V[\bar{n}_k] = \pi k^2. \quad (3.3.3)$$

In particular  $\bar{n}_0$  is a classical vacuum. Due to SO(3) symmetry the vacuum valley is isomorphic to  $S^2$ .

Now we turn to the sphalerons. For this we have to determine the Hessian of the energy functional, eq. (3.3.1), at the field  $\bar{n}_k$ , eq. (3.3.2). This is easy in spherical coordinates where eq. (3.3.2) reads  $\vartheta_k(x_2) = \frac{\pi}{2}$ ,  $\varphi_k(x_2) = kx_2$  and the metric is given by  $(g_{ij}) = \text{diag}(1, \sin^2 \vartheta)$ . Substituting these formulas in the static version of eq. (3.2.4) we obtain

$$\mathcal{H}[\bar{n}_k] = \frac{1}{2} \begin{pmatrix} -\partial_2^2 - k^2 & 0 \\ 0 & -\partial_2^2 \end{pmatrix}. \quad (3.3.4)$$

One immediately sees that  $k = \pm 1$  is the only solution with exactly one unstable mode ( $\delta\vartheta(x_2) = 1$ ,  $\delta\varphi(x_2) = 0$ ). So<sup>1</sup> the sphaleron solutions are given by  $\bar{n}_1$  and

<sup>1</sup>At this point we do not prove that these saddle points are true sphalerons in the mini-max sense. This proof can be constructed easily with the results of the next section.

SO(3) transformations thereof. In particular  $k \rightarrow -k$  can be undone by a rotation over  $\pi$  around any vector  $\vec{n}_k(x_2)$  ( $x_2$  fixed). Thus, sphaleron moduli-space is isomorphic to SO(3).

One can verify that SO(3) rotations are responsible for the 3 zero-modes of the Hessian. Note that the sphaleron is invariant under an  $x_2$ -translation in combination with a specific SO(2)  $\subset$  SO(3) rotation (in the case of eq. (3.3.2) around the 3-axis). Therefore spatial translations do not give new sphaleron solutions. Also one can check that the discrete symmetries  $x_2 \rightarrow -x_2$  and  $\vec{n} \rightarrow -\vec{n}$  leave invariant the sphaleron moduli-space.

## 3.4 Instanton solutions

### 3.4.1 Construction of the solutions

In stereographic projection the instanton equation (3.2.8) is particularly simple [64]:

$$\partial_z u = 0. \quad (3.4.1)$$

Here we have introduced complex coordinates on space-time,  $z = x_1 + ix_2$ ,  $\bar{z} = x_1 - ix_2$ , with derivatives  $\partial_z = \frac{1}{2}(\partial_1 - i\partial_2)$ ,  $\partial_{\bar{z}} = \frac{1}{2}(\partial_1 + i\partial_2)$ . The construction of instantons reduces to finding all analytic functions on  $T^1 \times \mathbb{R}$  with topological charge  $Q = 1$ .

Substitution of eq. (3.2.3) in eq. (3.2.6) gives, for any  $u$  satisfying eq. (3.4.1),

$$Q[u] = \frac{1}{\pi} \int_{T^1 \times \mathbb{R}} d^2x \frac{|\partial_z u|^2}{(1 + |u|^2)^2} = -\frac{1}{\pi} \int_{T^1 \times \mathbb{R}} d^2x \partial_{\bar{z}} \left( \frac{1}{1 + |u|^2} \frac{\partial_z u}{u} \right). \quad (3.4.2)$$

Handling carefully the poles of  $\frac{1}{1 + |u|^2} \frac{\partial_z u}{u}$ , i.e. the zeros of  $u$ , one finds

$$Q[u] = -\frac{1}{2\pi} \left( \int_{x_1=\infty} - \int_{x_1=-\infty} \right) dx_2 \frac{1}{1 + |u|^2} \frac{\partial_z u}{u} + \sum_i n_i, \quad (3.4.3)$$

where the  $x_2$  integrations run from 0 to  $2\pi$  and  $i$  runs over the zeros of  $u$ , of degree  $n_i \in \mathbb{N}$ . In order to simplify this formula we observe that both the kinetic and the potential term in the Lagrangian are semi-positive definite. Hence any finite-action configuration, in particular an instanton, must approach a point in the vacuum valley for  $x_1 \rightarrow \pm\infty$ :  $\lim_{x_1 \rightarrow \pm\infty} u(z) = u_{\pm}$ . In the derivation below we will assume that  $0 < |u_{\pm}| < \infty$ , as can always be achieved by an SO(3) rotation<sup>2</sup>. Under this assumption eq. (3.4.3) reduces to  $Q[u] = \sum_i n_i$ .

Now we are fully prepared to determine all instanton solutions. The strategy is first to find a class of solutions and then to prove no solutions exist outside this class. In order to construct a solution observe that

<sup>2</sup>This is equivalent to changing coordinates on  $S^2$  by shifting the pole used in the stereographic projection.

1. Since  $e^{z+2\pi i} = e^z$ , any  $u = h(e^z)$  ( $h$  single-valued and analytic) is a function on  $T^1 \times \mathbb{R}$  satisfying the instanton equation.
2. Under the above assumption any instanton can have only one zero  $z_1$  of degree  $n_1 = 1$ .

A class of functions satisfying all requirements is

$$u_{a,b,c,d}^{\text{inst}}(z) = -\frac{c + de^z}{a + be^z}, \quad (3.4.4)$$

with certain restrictions on the complex coefficients  $a, b, c, d$ . Note that  $u_+ = -d/b$ ,  $u_- = -c/a$  and  $z_1 = \ln(-c/d)$  (which is unique on  $T^1 \times \mathbb{R}$ ), so in order to satisfy the assumption  $0 < |u_{\pm}| < \infty$  we must require  $a, b, c, d \neq 0$ . Also we must demand  $a/b \neq c/d$  as otherwise the zero of  $u_{a,b,c,d}^{\text{inst}}$  cancels against its pole and we have a trivial solution with  $Q = 0$ .

To prove that all instanton solutions are of the form (3.4.4) is easy: suppose  $\bar{u}$  is an instanton. We can still assume  $0 < |\bar{u}_{\pm}| < \infty$ , so  $\bar{u}$  can have only one zero of degree 1, say at  $z = \bar{z}_1$ . Now define a function  $f$  by

$$\bar{u}(z) = u_{a,b,c,d}^{\text{inst}}(z)f(z). \quad (3.4.5)$$

By choosing  $-c/a = \bar{u}_-$ ,  $-d/b = \bar{u}_+$ ,  $c/d = -e^{\bar{z}_1}$  and imposing the instanton requirements  $\partial_{\bar{z}}\bar{u} = 0$ ,  $Q[\bar{u}] = 1$  and  $\bar{u}(z + 2\pi i) = \bar{u}(z)$ , that are already met by  $u$ , we see that  $f$  has to satisfy

$$\left\{ \begin{array}{l} \partial_{\bar{z}}f = 0 \\ \lim_{\text{Re}(z) \rightarrow \pm\infty} f(z) = 1 \\ f(z + 2\pi i) = f(z) \\ f(z) \neq 0. \end{array} \right. \quad (3.4.6)$$

Thus,  $1/f$  is analytic and bounded on  $\mathbb{C}$ . By Liouville's theorem this implies, using the second condition in eq. (3.4.6), that  $f(z) = 1$ . This completes the proof.

Finally we drop the assumption  $0 < |u_{\pm}| < \infty$ . Taking into account the boundary terms in eq. (3.4.3), one finds that the only restriction on  $(a, b, c, d)$  is that  $u_{a,b,c,d}^{\text{inst}}(z)$  is not constant, corresponding to

$$ad - bc \neq 0. \quad (3.4.7)$$

Hence the class of instantons, eq. (3.4.4), is precisely the set of conformal mappings of  $e^z$ .

By the same line of argument each multi-instanton with topological charge  $Q$  can be written as  $\prod_{n=1}^Q u_{a_n, b_n, c_n, d_n}^{\text{inst}}$ . To obtain (multi-)anti-instantons one substitutes  $z \rightarrow \bar{z}$ . These solutions are very similar to the well-known instantons on a space-time  $\mathbb{R}^2$  [64] (for details see the treatment of the equivalent  $CP^1$  model in ref. [71], or ref. [72] for a derivation directly on  $S^2$ ). The only difference is that  $z$  is substituted by  $e^z$ . On a space-time  $T^2$  the instanton formulas are rather more complicated [66].

In the context of an infinite space ( $L = \infty$ ) at non-zero temperature (i.e. finite  $T = (\text{temp.})^{-1}$ ) the solutions (3.4.4) were obtained before in ref. [68].

### 3.4.2 Physical interpretation of the moduli space

It is clear that the moduli space has six real dimensions:  $(a, b, c, d) \in \mathbb{C}^4$  with the projective character  $u_{ga,gb,gc,gd}^{\text{inst}} = u_{a,b,c,d}^{\text{inst}}$  for any  $g \in \mathbb{C} \setminus \{0\}$  (while the set of coefficients not satisfying eq. (3.4.7) has zero measure). A physical parametrization of this moduli space is  $(-c/a, -d/b, \ln(-c/d))$ , corresponding to the starting point  $u_-$  in the vacuum valley  $S^2$ , the end point  $u_+$ , and a space-time translation parameter that is easily shown to be in 1-1 relation with the instanton position (cf. eq. (3.4.10) below). The disadvantage of this parametrization is that it leaves unclear what kind of manifold the moduli space is. The parametrization is singular for  $u_- = u_+$  since in this case the requirement (3.4.7) is not met. Note that this means that even for  $T \rightarrow \infty$  periodic boundary conditions do not admit instanton solutions. We will see below that  $u_- \rightarrow u_+$  corresponds to  $\rho \rightarrow 0$ , where  $\rho$  is the instanton scale parameter.

For a better description of moduli space it is necessary first to consider the transformation of instantons under space-time translations and  $\text{SO}(3)$  rotations. We will see that most instantons are not invariant under these symmetries which therefore give rise to five of the six dimensions of moduli space. The interesting sixth parameter, not related to a symmetry of the action, will play the role of a scale parameter, related to the geodesic distance between the points  $u_{\pm}$  in the vacuum valley  $S^2$ .

From eq. (3.4.4) it is trivial to see that under a translation  $z \rightarrow z + z_0$ ,  $(a, b, c, d) \rightarrow (a, be^{z_0}, c, de^{z_0})$ . Note that eq. (3.4.7) is therefore invariant under this shift, as it should be: any continuous symmetry of the action must be present in the instanton moduli-space.

The effect of an  $\text{SO}(3)$  rotation is more difficult to derive because while it acts linearly on  $\vec{n}$ ,  $\vec{\pi}$  and  $u$  are related non-linearly by eq. (3.2.3). Nevertheless, one can show that  $(a, b, c, d)$  again transform linearly. After some effort one sees that the rotation

$$\vec{n}(x) \rightarrow R\vec{n}(x), \quad R = e^{\alpha_a L^a}, \quad L^a_{ij} = -\varepsilon_{aij}, \quad (\alpha_a \in \mathbb{R}) \quad (3.4.8)$$

induces

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \rightarrow \tilde{R} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \quad \tilde{R} = e^{\alpha_a \tilde{L}^a},$$

$$\tilde{L}^a = -\frac{i}{2} \sigma^a \otimes \mathbf{1}, \quad (\text{i.e. } \tilde{L}^1 = -\frac{i}{2} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ etc.}), \quad (3.4.9)$$

where  $\sigma^a$  are the Pauli-matrices. Since  $\pm(a, b, c, d)$  are identified, this is a representation of  $\text{SO}(3)$ . Note that only  $(a, c)$  and  $(b, d)$  mix. Hence both  $|a|^2 + |c|^2$  and  $|b|^2 + |d|^2$  are rotationally invariant, as is eq. (3.4.7).

Using the projective character of moduli space, and an  $\text{SO}(3)$  rotation  $\tilde{R}$ , it is always possible to bring  $u_{a,b,c,d}^{\text{inst}}$  ( $ad \neq bc$ ) to the form  $u_{-1,0,\tilde{c},\tilde{d}}^{\text{inst}}$  with  $\tilde{c} \in \mathbb{R}$ ,  $\tilde{c} \geq 0$ ,  $\tilde{d} \in \mathbb{C} \setminus \{0\}$ . If  $\tilde{c} > 0$ , this fixes  $\tilde{R}$  completely. If  $\tilde{c} = 0$ ,  $\tilde{R}$  is only unique up to a factor

$e^{\alpha_3 \bar{L}^3}$ . This can be fixed by requiring  $\bar{d} = |\bar{d}|$ . Therefore we can parametrize  $u_{a,b,c,d}^{\text{inst}}$  uniquely by  $\bar{R}$ ,  $|\bar{d}|$  and  $\bar{c}e^{i\phi}$  ( $\bar{c} \geq 0$ ). Here  $\phi = \text{Arg}(\bar{d})$  ( $\phi \in [0, 2\pi)$ ) if  $\bar{c} > 0$  and  $\phi$  is undetermined if  $\bar{c} = 0$ . We conclude that  $(\bar{c}, \phi)$  can be viewed as polar coordinates on  $\mathbb{R}^2$ . Furthermore,  $|\bar{d}| > 0$  due to eq. (3.4.7), so  $\ln|\bar{d}| \in \mathbb{R}$ . Thus the instanton moduli space is isomorphic to  $\text{SO}(3) \times \mathbb{R}^2 \times \mathbb{R}$ . The discrete symmetry transformations  $\bar{n} \rightarrow -\bar{n}$  (in stereographic coordinates  $u \rightarrow -1/\bar{u}$ ),  $x_2 \rightarrow -x_2$  or  $x_1 \rightarrow -x_1$  make an instanton solution anti-analytic, and therefore are transformations from instanton moduli-space into anti-instanton moduli-space. This should be compared to the sphaleron moduli-space which is invariant under such discrete transformations.

Note that  $\bar{d}e^z = e^{z + \ln|\bar{d}| + i\phi}$ . So after an  $\text{SO}(3)$  rotation and a translation in space-time  $T^1 \times \mathbb{R}$ , any instanton solution can be brought to the form

$$u_c^{\text{inst}}(z) \equiv c + e^{z + \ln(\sqrt{1+c^2}) + i\pi}, \quad (c \geq 0). \quad (3.4.10)$$

The factor  $e^{\ln(\sqrt{1+c^2}) + i\pi} = -\sqrt{1+c^2}$  centers the instanton around  $z = 0$  (see below). From the above equation it follows that  $\lim_{x_1 \rightarrow -\infty} |u_c^{\text{inst}}(z)| = \infty$  and  $\lim_{x_1 \rightarrow -\infty} u_c^{\text{inst}}(z) = c$ , hence (using eq. (3.2.3)) this instanton 'tunnels' from  $(\sin \vartheta_-, 0, \cos \vartheta_-)$  to  $(0, 0, 1)$ , with  $\cot \frac{1}{2}\vartheta_- = c$ . Note that  $\theta_-$  is the geodesic distance between these vacua.

Let us determine the instanton size  $\rho$  as function of  $c = \cot \frac{1}{2}\vartheta_-$ . Substituting eq. (3.4.10) into the Lagrangian density, which for an instanton in stereographic coordinates is  $4\pi$  times the integrand in eq. (3.4.2) (see eqs. (3.2.7), (3.2.8)), one obtains

$$\mathcal{L}[u_c^{\text{inst}}](x_1, x_2) = 4 \frac{|\partial_2 u_c^{\text{inst}}|^2}{(1 + |u_c^{\text{inst}}|^2)^2} = \frac{1}{(\sqrt{1+c^2} \cosh x_1 - c \cos x_2)^2}. \quad (3.4.11)$$

This function is plotted in fig. 3-1 for different values of  $c$ . From the formula it is clear that  $u_c^{\text{inst}}$  is centered at  $z = 0$ . Now consider the potential along the instanton path,

$$V[u_c^{\text{inst}}](x_1) = \frac{1}{2} \int_0^{2\pi} dx_2 \mathcal{L}[u_c^{\text{inst}}](x_1, x_2). \quad (3.4.12)$$

The factor  $\frac{1}{2}$  comes from the fact that the kinetic energy,  $\frac{1}{2} \int_{T^1} dx_2 g_{ij}(v) \partial_1 v^i \partial_1 v^j$ , is equal to the potential energy, eq. (3.3.1) (this 'self-duality' follows from eq. (3.2.8)). We see that the potential is maximal at  $x_1 = 0$  where it satisfies

$$V_c^{\text{max}} \equiv V[u_c^{\text{inst}}](0) = \pi \sqrt{1+c^2}. \quad (3.4.13)$$

Since all instantons have equal action, it is natural to define the instanton size

$$\rho(c) \equiv \frac{\pi}{V_c^{\text{max}}} = \frac{1}{\sqrt{1+c^2}}. \quad (3.4.14)$$

Note however that for small  $c$  there are two different scales: from fig. 3-1 we see that the shape of  $\mathcal{L}[u_c^{\text{inst}}](x_1, x_2)$  is anisotropic. Only for  $c \gg 1$  and  $x_1^2 + x_2^2 \ll 1$  the boundary effects disappear and  $\mathcal{L}[u_c^{\text{inst}}](x_1, x_2) \approx \frac{4c^2}{(1+c^2(x_1^2+x_2^2))^2}$  becomes rotationally invariant.

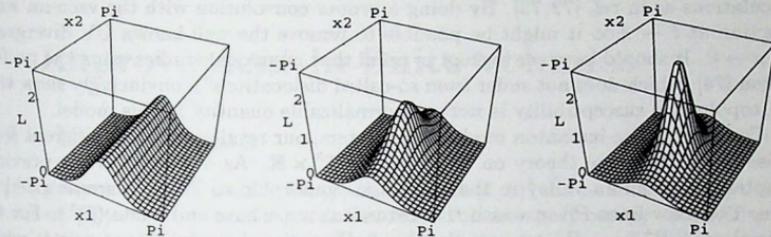


Figure 3-1. The Lagrangian densities of three instantons, eq. (3.4.10), with from left to right  $c = 0; 0.25; 0.5$ .

The relation between instantons and sphalerons is now also clear. From eqs. (3.4.13), (3.3.3) we see that only  $u_{c=0}^{\text{inst}}$  can go through a sphaleron at the time of maximal  $V[u_c^{\text{inst}}]$  (i.e.  $x_1 = 0$ ). Indeed this does happen, since eq. (3.4.10) gives  $u_{c=0}^{\text{inst}}(x_1 = 0, x_2) = -e^{ix_2}$ , which up to a rotation is precisely the sphaleron solution  $\vec{n}_1$ , eq. (3.3.2), in stereographic coordinates. It is easy to show that  $u_{c=0}^{\text{inst}}$  as a function of  $x_1$  is a stream line of the energy functional. A strongly related observation is that  $\partial_1 u_{c=0}^{\text{inst}}|_{x_1=0} = -e^{ix_2}$  corresponds to the unstable mode of the sphaleron solution.

Finally note that  $u_{c=0}^{\text{inst}}$ , unlike  $u_{c>0}^{\text{inst}}$ , is a point in instanton moduli-space that is symmetric under a joint  $\text{SO}(2)$  rotation  $e^{\alpha_3 \tilde{L}^3}$  and a spatial translation  $x_2 \rightarrow x_2 - \alpha_3$ . The sphaleron has of course the same symmetry, as mentioned in section 3.3. Also a natural correspondence between sphaleron moduli-space  $\text{SO}(3)$  and the subspace of widest instantons emerges. From the paragraph above eq. (3.4.10) it follows that the latter is isomorphic to  $\text{SO}(3) \times \mathbb{R}$ ,  $\mathbb{R}$  corresponding to time translations.

### 3.5 Conclusion

We have proven that the  $\text{O}(3)$   $\sigma$ -model on a space-time  $\mathbb{T}^1 \times \mathbb{R}$  admits instantons. The moduli space is 6-dimensional ( $\text{SO}(3) \times \mathbb{R}^2 \times \mathbb{R}$ ): 3 parameters for  $\text{SO}(3)$ , 2 for scaling and spatial translations, 1 for time translations. It is possible to 'tunnel' between any different points  $\vec{n}_\pm$  in the vacuum valley ( $S^2$ ), but this gives no independent parameters. Three parameters describing  $\vec{n}_\pm$  can be removed by an  $\text{SO}(3)$  rotation, while the fourth (the geodesic distance between the points) depends uniquely on the scale parameter  $\rho$ . Instantons with maximum scale, as set by the extent of spatial  $\mathbb{T}^1$ , satisfy  $\vec{n}_+ = -\vec{n}_-$ . These, and only these, instantons go through sphalerons. Exotic possibilities like the one sketched in the introduction do not take place in this model. On the other hand,  $\rho \rightarrow 0$  corresponds to  $\vec{n}_+ \rightarrow \vec{n}_-$ . Exact equality cannot be reached.

We think this peculiar size dependence is important for improving on instanton gas

calculations as in ref. [72, 73]. By doing a proper convolution with the vacuum wave function at  $t \rightarrow \pm\infty$  it might be possible to remove the well-known UV divergence for  $\rho \rightarrow 0$ . It should however be kept in mind that numerical studies using the perfect action [74], which does not suffer from so-called dislocations<sup>3</sup>, convincingly show that the topological susceptibility is *not* a renormalizable quantity in this model.

Concerning the instanton moduli-parameters, our results are quite different from those in  $SU(2)$  gauge theory on a space-time  $T^3 \times \mathbb{R}$ . As discussed in the previous chapter, the vacuum valley in that model is isomorphic to  $T^3$  [8] (parametrized by three Polyakov lines  $P_i$ ) in which the instantons must have endpoints ( $P_i^\pm$ ). For the special case  $P_i^+ = -P_i^-$  we saw that an 8-dimensional moduli-space exists, which according to the argument in section 2.2 includes a scale parameter. On the other hand, for the  $O(3)$   $\sigma$ -model on a space-time  $T^1 \times \mathbb{R}$  we have just proven that anti-periodic boundary conditions,  $\bar{n}_+ = -\bar{n}_-$ , fix the instanton size<sup>4</sup>. In this sense the  $O(3)$   $\sigma$ -model on  $T^1 \times \mathbb{R}$  is an unsuitable toy model for  $SU(2)$  gauge theory on  $T^3 \times \mathbb{R}$ .

<sup>3</sup>La. 'instantons' at the scale of the lattice spacing with the same topological charge as large instantons, but with considerably lower action. They may thus give unphysical contributions to the path integral that survive in the continuum limit [38].

<sup>4</sup>We also note that no instantons exist with anti-periodic boundary conditions over a finite time  $T$ ,  $\bar{n}(x_2 + T, x_2) = -\bar{n}(x_1, x_2)$  (while the winding number  $Q$  is still a well-defined integer object). The reason is simple: in stereographic coordinates anti-periodic boundary conditions read  $u(z+T) = -1/\bar{u}(z)$ , which is incompatible with the instanton equation  $\partial_z u = 0$ .

## 4 Improvement in finite volumes

### 4.1 Introduction

$SU(N)$  gauge theories in small finite spatial volumes  $[0, L]^3$  can be treated analytically due to asymptotic freedom. In particular finite-volume spectroscopy in  $SU(2)$  pure gauge theory is theoretically well-understood, even up to intermediate volumes [8]. Here 'intermediate' refers to those values of  $L$  for which non-perturbative effects associated with the non-trivial classical vacuum structure<sup>1</sup> need to be taken into account, but the instanton and sphaleron configurations described in chapter 2 are not yet important. The simplest method used to push the calculation to intermediate volumes is a Born-Oppenheimer approximation, integrating out the fast modes in a perturbative way, while keeping the constant modes (in particular the vacuum-valley variables) as a background field, thus obtaining an effective action for the constant modes. The low-lying spectrum of the corresponding effective Hamiltonian is then computed non-perturbatively.

In the last paper in ref. [8] this method was used in the continuum formulation. In ref. [11] the calculation was extended to the lattice, using the standard Wilson action. For volumes well beyond the point where the non-perturbative effects due to the vacuum structure set in, the agreement with Monte Carlo data [12] is perfect. In the context of improved lattice actions we can now make a trivial but important observation: from the two calculations one knows the magnitude of the lattice artefacts carried by the Wilson action, and hence the room available for improvement. This makes the finite volume a natural setting to *analytically* study the effectiveness of improvement schemes.

In this chapter we partly extend the calculation in ref. [11] to Symanzik [24, 25] and tadpole [29] improved actions, which were introduced in chapter 1. For a technical reason, namely diagonalization of the Feynman propagator, we introduce a new Symanzik-improved action instead of using the standard Lüscher-Weisz [25] action. The use of the new action, which we call the square action, greatly simplifies the computation of the background-field effective action.

Nevertheless, even for the square action the extraction of the low-lying spectrum is severely hampered. The reason is that, as for any Symanzik-improved action, the square action includes Wilson loops extending over more than one lattice spacing. This introduces next to nearest neighbor couplings in the effective action, giving rise to modes that make the transfer matrix non-Hermitian [75]. Though these modes have unphysical energies of order  $1/a$  ( $a$  being the lattice spacing), they couple to the low-energy modes. This makes the construction of an effective Hamiltonian to

<sup>1</sup>In this chapter we consider only untwisted space-times  $T^3 \times \mathbb{R}$ , so that the vacuum valley is isomorphic to  $T^3$ , cf. eq. (1.1.4).

sufficient order in  $a$  problematic (for details see ref. [76]).

For this reason we here limit ourselves to a modest goal: the computation, for a Symanzik-improved action, of the Abelian part of the effective potential, i.e. the effective action (per unit of time) for static background fields in the classical vacuum valley (henceforth called the Abelian potential). We will be able to do this computation for the square action. By comparing to the Wilson and continuum cases we can therefore see how effective Symanzik improvement is (for this object), and how much room is left for tadpole improvement. However, it should be understood that the Abelian potential itself is *not* a spectral quantity. It is only one ingredient in the calculation of the low-lying spectrum. The last part of this chapter is dedicated to Monte Carlo simulations of the Wilson, Symanzik and tadpole improved actions.

The outline of this chapter is as follows. In section 4.2 we briefly explain the background field method. In section 4.3 we introduce the square action and from it calculate the Abelian potential, the improvement of which is discussed at the end of the section. Monte Carlo results are given in section 4.4 and their bearing on (tadpole) improvement is discussed in section 4.5.

Throughout this chapter we limit ourselves to  $SU(2)$ .

## 4.2 Background fields on the lattice

From chapter 2 we copy the Lüscher-Weisz action (2.4.8), but for later use we add a  $2 \times 2$  Wilson loop. Taking the liberty to deviate from the notation in ref. [25] we assign to this loop the coefficient  $c_4$ . The action and its expansion for small lattice spacings  $a$  then read (relating lattice and continuum fields by  $U_\mu(x) = P \exp(\int_0^a A_\mu(x + s\hat{\mu}) ds)$ )

$$S(\{c_i\}) \equiv \sum_x \text{Tr} \left\{ c_0 \sum_{\mu \neq \nu} \langle 1 - \nu \square_{x, \mu} \rangle + 2c_1 \sum_{\mu \neq \nu} \langle 1 - \nu \square_{x, \mu} \rangle + \frac{4}{3} c_2 \sum_{\mu \neq \nu \neq \lambda} \langle 1 - \nu \square_{x, \mu, \lambda} \rangle \right. \\ \left. + 4c_3 \sum_{\mu \neq \nu \neq \lambda} \langle 1 - \nu \square_{x, \mu, \lambda} \rangle + c_4 \sum_{\mu \neq \nu} \langle 1 - \nu \square_{x, \mu} \rangle \right\} \quad (4.2.1)$$

$$= -\frac{a^4}{2} (c_0 + 8c_1 + 8c_2 + 16c_3 + 16c_4) \sum_{x, \mu, \nu} \text{Tr} (F_{\mu\nu}(x))^2 \\ + \frac{a^6}{12} (c_0 + 20c_1 - 4c_2 + 4c_3 + 64c_4) \sum_{x, \mu, \nu} \text{Tr} (D_\mu F_{\mu\nu}(x))^2 \\ + \frac{a^6}{3} (c_2 + 3c_3) \sum_{x, \mu, \nu, \lambda} \text{Tr} (D_\mu F_{\mu\lambda}(x) D_\nu F_{\nu\lambda}(x)) \\ + \frac{a^6}{3} c_2 \sum_{x, \mu, \nu, \lambda} \text{Tr} (D_\mu F_{\nu\lambda}(x))^2 + \mathcal{O}(a^8). \quad (4.2.2)$$

In this chapter the domain of summation for  $x = (n_0 a, n_1 a, n_2 a, n_3 a)$  is  $n_0 \in \{0, 1, \dots, N_0 - 1\}$ ,  $n_i \in \{0, 1, \dots, N - 1\}$ . Here  $N \equiv L/a$  (not to be confused by the number of colors, which equals 2) and  $N_0 \equiv T/a$ ,  $T$  being the extent of the time direction (below

we will send  $T \rightarrow \infty$ . The normalization and tree-level improvement conditions given below eq. (2.4.8) generalize to

$$\begin{aligned} c_0 + 8c_1 + 8c_2 + 16c_3 + 16c_4 &= 1, \\ c_0 + 20c_1 - 4c_2 + 4c_3 + 64c_4 = c_2 = c_2 + 3c_3 &= 0. \end{aligned} \quad (4.2.3)$$

Again,  $c_2 + 3c_3 = 0$  does not apply on-shell (this is even valid at higher loop order, see ref. [25] or section 5.2). The choice of a convenient solution to the conditions is postponed to the next section. We use scale invariance to put  $L = 1$  in the remainder of this chapter, so that  $a = 1/N$ . The limit  $T \rightarrow \infty$  then translates into  $N_0 \rightarrow \infty$  (for  $N$  fixed), while the continuum limit is reached by taking  $N \rightarrow \infty$ .

The parametrization  $U_\mu(x) = P \exp(\int_0^a A_\mu(x + s\bar{\mu}) ds)$  is suitable for deriving the tree-level improvement conditions, but not for doing perturbation theory or a background field calculation. For perturbation theory it is much more convenient to use

$$U_\mu(x) = \exp(q_\mu(x)), \quad (4.2.4)$$

where  $q_\mu(x)$  lies in the Lie algebra  $\mathfrak{su}(2)$  (in the fundamental representation). In the background field approach one splits up  $q_\mu(x)$  in a background part and a fluctuation part, only integrating the latter in the path integral. We limit ourselves to static background fields in the vacuum valley, whose link variable will be denoted by  $\hat{U}_\mu(x)$ . In an appropriate gauge we have

$$\hat{U}_j(x) = \hat{U}_j = \exp(\frac{1}{2}iC_j\sigma_3/N), \quad \hat{U}_0(x) = 1, \quad (4.2.5)$$

where  $C_j \in \mathbb{R}$  is defined modulo  $4\pi$  (for additional comments see ref. [8]).

The split-up of  $q_\mu(x)$  is best implemented on the group level [11, 77],

$$U_\mu(x) = \exp(\hat{q}_\mu(x))\hat{U}_\mu, \quad (4.2.6)$$

and the Abelian potential  $V_1^{\text{ab}}(\vec{C})$  (up to an irrelevant constant) is then defined by a saddle point approximation around the background field:

$$\begin{aligned} \exp\left(-\frac{N_0}{N}V_1^{\text{ab}}(\vec{C})\right) &\equiv \int \mathcal{D}\vec{q} \det \mathcal{M}(\vec{C}, \vec{q}) \exp\left[-\frac{1}{g_0^2} \left(S(\vec{C}, \vec{q}) - \sum_{\vec{x}} \text{Tr} \mathcal{F}_{\text{gf}}^2(\vec{C}, \vec{q})\right)\right], \\ \int \mathcal{D}\vec{q} &\equiv \prod_{\vec{x}} \int_{\mathbb{R}^3} d^3\vec{q}(x), \\ \mathcal{M}(\vec{C}, \vec{q}) &\equiv \frac{\delta \mathcal{F}_{\text{gf}}(\vec{C}, \vec{q})}{\delta \Lambda}, \end{aligned} \quad (4.2.7)$$

for  $N_0 \rightarrow \infty$ . Some explanation is in order. We restrict ourselves to a one-loop computation, and in the background field context this means that a Gaussian saddle point approximation is sufficient. In particular we approximate the Haar measure by the Euclidean measure  $\mathcal{D}\vec{q}$  on the Lie algebra (with  $\vec{q}(x) = -\frac{1}{2}i \sum_a \sigma_a \vec{q}^a(x)$ ). To implement the boundary conditions only spatially periodic fields are included in the integration. For convenience we also take periodic boundary conditions in time.

Since  $\bar{q}(x)$  is integrated over the non-compact space  $\mathbb{R}^3$ , gauge fixing is necessary. To this purpose the term  $-\text{Tr} \mathcal{F}_{\text{gf}}^2(\bar{C}, \bar{q})$  is added to the action in eq. (4.2.7). Gauge invariance is ensured by the inclusion of the Faddeev-Popov determinant  $\mathcal{M}$ . In the definition of  $\mathcal{M}$ ,  $\Lambda(x)$  defines the gauge transformation  $U_\mu(x) \rightarrow \exp(\Lambda(x))U_\mu(x) \exp(-\Lambda(x + a\hat{\mu}))$  (see ref. [11] for details).

The symbol  $S(\bar{q}, \bar{C})$  appearing in eq. (4.2.7) is the action (4.2.1) with eq. (4.2.6) substituted. For the one-loop approximation it is of course sufficient to expand  $S(\bar{q}, \bar{C}) - \text{Tr} \mathcal{F}_{\text{gf}}^2(\bar{C}, \bar{q})$  to quadratic order in  $\bar{q}$ ,

$$S(\bar{q}, \bar{C}) - \sum_x \text{Tr} \mathcal{F}_{\text{gf}}^2(\bar{C}, \bar{q}) = - \sum_{x\mu\nu} \text{Tr} (\hat{q}_\mu \mathcal{W}_{\mu\nu}(\bar{C}) \hat{q}_\nu)(x) + \mathcal{O}(\bar{q}^4), \quad (4.2.8)$$

and to replace  $\det \mathcal{M}(\bar{C}, \bar{q})$  by  $\det \mathcal{M}(\bar{C}) \equiv \det \mathcal{M}(\bar{C}, 0)$ . Working out the integration then gives the formal result

$$\begin{aligned} \frac{1}{N} V_1^{\text{ab}}(\bar{C}) &= - \lim_{N_0 \rightarrow \infty} \frac{1}{N_0} \ln \frac{\det \mathcal{M}(\bar{C})}{\sqrt{\det \mathcal{W}(\bar{C})}} \\ &= - \lim_{N_0 \rightarrow \infty} \frac{1}{N_0} \left\{ \text{tr} \ln \mathcal{M}(\bar{C}) - \frac{1}{2} \text{tr} \ln \mathcal{W}(\bar{C}) \right\}, \end{aligned} \quad (4.2.9)$$

where 'det' and 'tr' run over the appropriate Hilbert spaces (associated with the vector field  $\hat{q}_\mu^\alpha(x)$  and the Faddeev-Popov ghost field  $\psi^a(x)$ , for  $\mathcal{W}$  and  $\mathcal{M}$  respectively). Note that the indices  $(\mu, a, x)$  take discrete values only.

### 4.3 Abelian potential for the square action

The expansion of  $S(\bar{C}, \bar{q})$  to second order in  $\bar{q}$  is not difficult. As an example, consider the contribution of the  $2 \times 1$  Wilson loop. Labeling the links by  $1 \dots 6$ , in counter-clockwise order starting at the point  $x$ , we have

$$\begin{aligned} \text{Diagram} &= U_1 U_2 U_3 U_4^\dagger U_5^\dagger U_6^\dagger = e^{q_1} (\hat{U}_1 e^{q_2} \hat{U}_1^\dagger) (\hat{U}_1 \hat{U}_2 e^{q_3} \hat{U}_2^\dagger \hat{U}_1^\dagger) \\ &\quad \times \hat{U}_1 \hat{U}_2 \hat{U}_3 \hat{U}_4^\dagger \hat{U}_5^\dagger \hat{U}_6^\dagger (\hat{U}_6 \hat{U}_5 e^{q_4} \hat{U}_5^\dagger \hat{U}_6^\dagger) (\hat{U}_6 e^{q_5} \hat{U}_6^\dagger)^\dagger e^{-q_6}. \end{aligned} \quad (4.3.1)$$

Using that  $\hat{U}_\mu$  is a constant Abelian field one then quickly derives

$$\begin{aligned} \text{Tr} \left( 1 - \text{Diagram} \right) &= \text{Tr} \left[ 1 - \exp(\hat{q}_\mu) \exp((1 + \hat{D}_\mu) \hat{q}_\mu) \exp((1 + \hat{D}_\mu)^2 \hat{q}_\mu) \right. \\ &\quad \times \exp(-(1 + \hat{D}_\nu)(1 + \hat{D}_\mu) \hat{q}_\mu) \exp(-(1 + \hat{D}_\nu) \hat{q}_\mu) \\ &\quad \left. \times \exp(-\hat{q}_\nu) \right] \\ &= -\frac{1}{2} \text{Tr} \left[ (2 + \hat{D}_\mu) (\hat{D}_\mu \hat{q}_\nu - \hat{D}_\nu \hat{q}_\mu) \right]^2 + \mathcal{O}(\hat{q}^4), \end{aligned} \quad (4.3.2)$$

where all terms are taken in the point  $x$  and

$$(\hat{D}_\mu \hat{q}_\nu)(x) \equiv \hat{U}_\mu \hat{q}_\nu(x + a\hat{\mu}) \hat{U}_\mu^\dagger - \hat{q}_\nu(x) \quad (4.3.3)$$

is the covariant background field lattice derivative (with  $\hat{\mu}$  the unit vector in the positive  $\mu$ -direction). Note that  $[\hat{D}_\mu, \hat{D}_\nu] = 0$  in the constant Abelian case.

For the other loops one uses the same method to obtain

$$S(\bar{C}, \bar{q}) = -\frac{1}{2} \sum_{x, \mu, \nu} \text{Tr} \left[ c_0 (\hat{D}_\mu \bar{q}_\nu - \hat{D}_\nu \bar{q}_\mu)^2 + 2c_1 \left\{ (2 + \hat{D}_\mu) (\hat{D}_\mu \bar{q}_\nu - \hat{D}_\nu \bar{q}_\mu) \right\}^2 + c_4 \left\{ (2 + \hat{D}_\nu) (2 + \hat{D}_\mu) (\hat{D}_\mu \bar{q}_\nu - \hat{D}_\nu \bar{q}_\mu) \right\}^2 \right] (x) + \mathcal{O}(\bar{q}^4). \quad (4.3.4)$$

Here we chose  $c_2 = c_3 = 0$ , which is consistent with the improvement conditions (4.2.3), because the corresponding loops are not needed in the construction of the square action. At tree level this construction simply amounts to choosing

$$c_4 = z^2 c_0, \quad z \equiv c_1 / c_0. \quad (4.3.5)$$

In that case one can complete squares (hence the name of our action) and write

$$S(\bar{C}, \bar{q}) = -\frac{1}{2} c_0 \sum_{x, \mu, \nu} \text{Tr} \left[ (1 + z(2 + \hat{D}_\mu^\dagger)(2 + \hat{D}_\mu)) (\hat{D}_\mu \bar{q}_\nu - \hat{D}_\nu \bar{q}_\mu) \times (1 + z(2 + \hat{D}_\nu^\dagger)(2 + \hat{D}_\nu)) (\hat{D}_\mu \bar{q}_\nu - \hat{D}_\nu \bar{q}_\mu) \right] + \mathcal{O}(\bar{q}^4), \quad (4.3.6)$$

where the Hermitian conjugation  $\dagger$  is defined with respect to  $\sum_x \text{Tr}$ . The reason why this is useful, is that we can now diagonalize the tensor  $\mathcal{W}_{\mu\nu}(\bar{C})$ , see eq. (4.2.8), with respect to the Lorentz structure by choosing a (modified) covariant gauge<sup>2</sup>:

$$\mathcal{F}_{\text{gr}}(\bar{C}, \bar{q})(x) = \sqrt{c_0} \sum_{\mu} (1 + z(2 + \hat{D}_\mu^\dagger)(2 + \hat{D}_\mu)) \hat{D}_\mu^\dagger \bar{q}_\mu(x). \quad (4.3.7)$$

Explicitly the result for  $\mathcal{W}_{\mu\nu}$  reads

$$\mathcal{W}_{\mu\nu} = c_0 \delta_{\mu\nu} \sum_{\lambda} (1 + z(2 + \hat{D}_\mu^\dagger)(2 + \hat{D}_\mu)) (1 + z(2 + \hat{D}_\lambda^\dagger)(2 + \hat{D}_\lambda)) \hat{D}_\lambda^\dagger \hat{D}_\lambda. \quad (4.3.8)$$

One can check that the Faddeev-Popov matrix equals (cf. ref. [11])

$$\mathcal{M} = \sqrt{c_0} \sum_{\lambda} (1 + z(2 + \hat{D}_\lambda^\dagger)(2 + \hat{D}_\lambda)) \hat{D}_\lambda^\dagger \hat{D}_\lambda. \quad (4.3.9)$$

The above diagonalization of  $\mathcal{W}_{\mu\nu}$  with respect to  $(\mu, \nu)$  allows one to compute the Abelian potential in technically much the same way as for the Wilson action [11].

In the above  $z$  is a free parameter. For  $z = 0$  one recovers the Wilson action. On the other hand the action is improved for

$$c_0 = \frac{16}{9}, \quad c_1 = -\frac{1}{9}, \quad c_2 = 0, \quad c_3 = 0, \quad c_4 = \frac{1}{144}, \quad (4.3.10)$$

which clearly satisfies eqs. (4.2.3) and (4.3.5) (with  $z = -\frac{1}{16}$ ). The action (4.2.1) for this set of coefficients is what we call the square action. Note that the diagonalization

<sup>2</sup>This is similar to the modified Coulomb gauge introduced in ref. [27].

procedure is not possible for the standard Lüscher-Weisz improved action [25], i.e.  $c_0 = \frac{5}{3}$ ,  $c_1 = -\frac{1}{12}$  and  $c_2 = c_3 = c_4 = 0$ . Also note that tadpole improvement [29], where one replaces  $U_\mu(x) \rightarrow U_\mu(x)/u_0$ , can be naturally embedded in our approach because it leaves invariant the diagonalization condition  $c_0 c_4 = c_1^2$ . For the tadpole-improved square action one easily finds  $z = -1/(16u_0^2)$  and  $c_0 = 1/(1+4z)^2$ .

To compute the determinants in eq. (4.2.9) we first complete the diagonalization of  $\mathcal{W}$  and  $\mathcal{M}$ . Diagonalization with respect to space-time coordinates is trivial by a Fourier transformation. Diagonalization with respect to  $SU(2)$  amounts to finding the eigenvalues of the conjugation with  $\hat{U}_\mu$ , eq. (4.2.5). The eigenvectors are simply  $\sigma_3$  and  $\sigma_\pm \equiv \sigma_1 \pm \sigma_2$ , with eigenvalues  $\exp(isC_\mu/N)|_{s=0,\pm 1}$  (where  $C_0 = 0$ ). Hence  $\hat{D}_\mu$  has eigenvalues

$$\begin{aligned} \bar{D}_\mu(k, s) &= \exp(ik_\mu^s) - 1, \\ k_\mu^s &\equiv k_\mu + sC_\mu/N, \quad k_\mu \equiv 2\pi n_\mu/N_\mu, \quad (N_j \equiv N), \\ s &\in \{-1, 0, 1\}, \quad n_0 \in \{0, 1, \dots, N_0 - 1\}, \quad n_j \in \{0, 1, \dots, N - 1\}. \end{aligned} \quad (4.3.11)$$

Note that  $\exp(ik_\mu^s) - 1$  is the Fourier transform of the conventional lattice derivative  $(\partial_\mu \varphi)(x) \equiv \varphi(x + a\hat{\mu}) - \varphi(x)$ .

The eigenvalues of  $\mathcal{W}$  and  $\mathcal{M}$  thus read (with  $\mathcal{W}_{\mu\nu} = \mathcal{W}_\mu \delta_{\mu\nu}$ )

$$\begin{aligned} \mathcal{M}(\bar{C}, s, k) &= \sqrt{c_0} \sum_\rho \omega^2(k_\rho^s), \\ \mathcal{W}_\mu(\bar{C}, s, k) &= \sqrt{c_0} \lambda(k_\mu^s) \mathcal{M}(\bar{C}, s, k), \\ \omega^2(y) &\equiv 4 \sin^2(\frac{1}{2}y) \lambda(y), \\ \lambda(y) &\equiv 1 + 4z \cos^2(\frac{1}{2}y). \end{aligned} \quad (4.3.12)$$

Substitution in eq. (4.2.9) gives (up to an irrelevant  $\bar{C}$ -independent term)

$$\begin{aligned} V_1^{\text{ab}}(\bar{C}) &= N \sum_{n,s} \left\{ \frac{1}{2} \left( \sum_i \ln \lambda(k_i^s) \right) + \lim_{N_0 \rightarrow \infty} \frac{1}{N_0} \sum_{n_0} \left[ \ln \left( \cosh \Omega_+(k^s) - \cos k_0 \right) \right. \right. \\ &\quad \left. \left. + \ln \left( \cosh \Omega_-(k^s) - \cos k_0 \right) \right] \right\}, \end{aligned} \quad (4.3.13)$$

where  $\cosh \Omega_\pm(k^s)$  are the algebraic roots of  $\mathcal{M}(\bar{C}, s, k)$  with respect to  $\cos k_0$ . These are easily found to be (with  $\omega^2(k) \equiv \sum_i \omega^2(k_i)$ )

$$\cosh \Omega_\pm(k) = -\frac{1}{4z} \left( 1 \pm \sqrt{(1+4z)^2 + 4z\omega^2(k)} \right). \quad (4.3.14)$$

Their contributions to eq. (4.3.13) are both of the Wilson form<sup>3</sup>, so the result of ref. [11] can easily be generalized to the square action. For the sake of completeness we give a short derivation.

<sup>3</sup>For  $z \rightarrow 0$ ,  $\ln \lambda(k_i) \rightarrow 0$  and  $\Omega_\pm(k^\pm)$  becomes  $\bar{C}$  independent, so only  $\Omega_-(\bar{C}, k^\pm) \rightarrow \omega(k^\pm)$  contributes to the  $\bar{C}$  dependence of  $V_1^{\text{ab}}(\bar{C})$ .

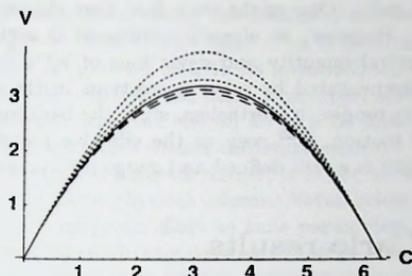


Figure 4-1. The effective potential  $V = V_1^{\text{ab}}$  for a constant Abelian background field  $A_1 = \frac{1}{2}iC\sigma_3/N$ . The full line represents the continuum result (obtained by taking the number of lattice spacings  $N \rightarrow \infty$ ). The lower two dashed curves are for the square action with  $N = 3$  and  $4$  ( $N = 6$  is indistinguishable from the continuum curve). The upper three dotted curves are for the Wilson action with  $N = 3, 4$  and  $6$ .

Due to the limit  $N_0 \rightarrow \infty$ , the sum over  $n_0$  in eq. (4.3.13) reduces to an integral, which due to the periodicity with respect to  $k_0 \in [0, 2\pi]$  may be written as a contour integral in the variable  $w \equiv \exp(ik_0)$ . From the residue theorem one then immediately derives (for  $\Omega > 0$ )

$$\frac{\partial}{\partial(\cosh \Omega)} \left( \lim_{N_0 \rightarrow \infty} \frac{1}{N_0} \ln(\cosh \Omega - \cos k_0) \right) = \frac{1}{\sinh \Omega}, \quad (4.3.15)$$

because  $w = \exp(-\Omega)$  is the only pole inside the contour  $|w| = 1$ . Hence we see that (again up to a  $\vec{C}$ -independent term)

$$\begin{aligned} V_1^{\text{ab}}(\vec{C}) &= N \sum_{\mathbf{n}, s} \left\{ \Omega_+(\mathbf{k}^s) + \Omega_-(\mathbf{k}^s) + \frac{1}{2} \sum_i \ln \lambda(k_i^s) \right\} \\ &= N \sum_{\mathbf{n}} \left\{ 4 \operatorname{asinh} \left( \frac{1}{\sqrt{-4z}} \sqrt{1+4z + \frac{1}{2}\omega^2 + \omega\sqrt{1+\frac{1}{4}\omega^2}} \right) + \sum_i \ln \lambda(k_i^+) \right\}, \end{aligned} \quad (4.3.16)$$

where the second step requires some rearranging, and  $\omega$  is shorthand for  $\omega(\mathbf{k}^+)$ . Due to the summation over  $\mathbf{n}$  and the fact that  $\omega(y)$  and  $\lambda(y)$  are even, the  $s = \pm 1$  contributions to  $V_1^{\text{ab}}(\vec{C})$  are equal. Furthermore, the  $s = 0$  contribution is independent of  $\vec{C}$ . Eq. (4.3.16) is our final result. Note that the sum over  $\mathbf{n}$  contains only  $N^3$  terms and can be computed numerically up to large values of  $N = L/a$ .

In fig. 4-1  $V_1^{\text{ab}}(C, 0, 0)$  is plotted for the Wilson ( $z = 0$ ) and square ( $z = -\frac{1}{16}$ , i.e.  $u_0 = 1$ ) actions, for small values of  $N$ . Also the continuum curve ( $N = \infty$ ) is included. For the square action the result for  $N = 6$  can already not be distinguished from that in the continuum at the scale of this figure. For  $N = 3, 4$  we see that the square action improves the Wilson action to a large extent, so that tadpole improvement cannot

give much additional gain. One might even fear that choosing  $u_0 \neq 1$  will make the agreement worse. However, as already mentioned in section 4.1, the effective potential is not a spectral quantity and deviations of  $V_1^{ab}(\bar{C})$  from the continuum can in principle be compensated by other corrections in the effective Hamiltonian for the zero-momentum modes. Nevertheless, since the background field satisfies the (lattice) equations of motion,  $V_1^{ab}$  seen as the effective potential obtained from a Legendre transformation is a well defined and gauge invariant object.

## 4.4 Monte Carlo results

By means of Monte Carlo simulations we have determined the masses of the scalar ( $A_1^+$ ) and tensor glueballs in intermediate volumes. Due to the breaking of rotational invariance, the tensor glueballs split in the doublet  $E^+$  and the triplet  $T_2^+$ . Also the energies of the electric flux states (called torelons in ref. [12]) with one, two and three units of electric flux ( $e_i$ ,  $i = 1, 2, 3$ ) were measured. In addition we considered the states with two ( $T_{11}^+$  or  $B(110)$ ) and three ( $T_2^+(111)$ ) units of electric flux that have  $T_2^+$  quantum numbers (negative parity in two directions of electric flux, symmetrized in those two directions). See ref. [12] for details and further references.

The size of the lattice used was  $4^3 \times 128$ . Masses  $m_i$  are converted to dimensionless parameters  $z_i \equiv m_i L$ . In lattice units we hence multiply masses with the number of lattice sites in the spatial directions. In large volumes one should have  $z_{e_k} = \sigma L^2 \sqrt{k}$ , where  $\sigma$  is the infinite volume string tension. This is why we will consider the ratios  $\sqrt{z_{e_i}}/z_{A_1^+}$ . These and other mass ratios will be plotted as a function of  $z_{A_1^+}$ . The analytic result for the Wilson action [11], derived by diagonalizing the effective Hamiltonian to describe low-lying states, is valid up to  $z_{A_1^+} \approx 5$ , after which degrees of freedom that were integrated out perturbatively will receive non-perturbative contributions from transitions over the sphaleron determined in chapter 2. The breakdown will occur at smaller volumes for higher excited states. For the Wilson action we have chosen  $\beta \equiv 4/g_0^2 = 3.0$  and  $\beta = 2.4$ ; for the improved actions  $\beta$  was tuned to yield results in roughly the same physical volume. These parameters correspond to lattice spacings of approximately  $a = 0.018$  fermi and  $a = 0.12$  fermi, if one uses the string tension to set the scale. For the smallest of the two, one expects tree-level improvement to be effective and we have therefore not tadpole improved the actions in this case. Note that for these small volumes one finds from the analytic results that the lattice artefacts in the mass ratios are quite much bigger [11] than in larger volumes. Data was taken for both the Lüscher-Weisz (LW) and square Symanzik actions, and as a test on our programs also for the Wilson action for which we can compare with available high-precision data [12].

At the larger volume we concentrated our attention to the Lüscher-Weisz action with tree-level and (following the prescription<sup>4</sup> of refs. [29,31]) tadpole improved one-

<sup>4</sup>The general prescription for tadpole improvement would be to replace the Symanzik coefficient functions  $c_i(g_0^2)$  by  $\tilde{c}_i(g_0^2)/u_0^{n_i}$ , where  $n_i + 4$  equals the number of links in the  $i^{\text{th}}$  kind of Wilson loop

loop [26] values of the coefficients. Using the results of the next chapter one could of course also use the square action. The value of  $u_0$  is determined self consistently, adjusting the input value of  $u_0$  to agree with its measured value. This only requires little Monte Carlo time. We verified that there is no observable volume dependence of  $u_0$  by comparing its value with the one on an  $8^3 \times 64$  lattice (the difference was less than 0.3%, consistent with zero within statistical errors).

Because of the availability of analytic results, it is not necessary to exactly tune the different actions to the same physical volume. Nevertheless, in particular for the data at  $a = 0.12$  fermi we made an effort to tune parameters appropriately, as we can make a stronger point when directly comparing lattice data at the same physical volume. The results are presented in fig. 4-2, together with the analytic results for the continuum (solid curves) and for the Wilson action on a lattice of size  $4^3 \times \infty$ .

## 4.5 Discussion

As was to be expected, at lattice spacings of approximately 0.02 fermi ( $z_{A_4^+} \approx 2$ ), tree-level improvement is seen to bring the lattice results quite close to those of the continuum, both for the Lüscher-Weisz and square actions. In both cases the improvement is considerable.

Also at lattice spacings around 0.12 fermi and volumes of approximately 0.48 fermi ( $z_{A_4^+} \approx 4$ ), the agreement of the Wilson action lattice data with the corresponding analytic results is in general very good for the lowest-lying states. Any deviations, at least when they are larger than the statistical errors, should be due to additional non-perturbative corrections caused by transitions over the sphaleron, not included in the calculation in ref. [11]. We expect them to be roughly the same for the Wilson and continuum actions. Therefore the difference in the analytic results between the continuum and Wilson actions can be used to estimate how far the improved data is removed from the continuum values. Significant improvement is observed in some of the cases, in particular for  $z_{T_1^+}/z_{A_4^+}$ , approaching the continuum analytic result.

The most salient feature of our data is that, for the ratios considered, tadpole improvement has no significant additional effect compared to Symanzik improvement. Perhaps for the cases where Symanzik improvement is already significant this is what one would want, but our results show some instances where tree-level Symanzik improvement has no effect and tadpole improvement is of no help either.

In particular we note that the ratio  $\sqrt{z_{T_1^+}}/z_{A_4^+}$ , measured to an accuracy better than 1.5%, deviates from its continuum value by 5-6%. For this quantity tree-level Symanzik improvement as well as tadpole improvement is unable to show deviations from the Wilson result. This puts some doubt on the usefulness of tadpole improvement for careful extrapolations of mass ratios to the continuum limit.

(cf. eq. (1.2.17)). To one-loop order, the function  $c_1(g_0^2)$  is determined by  $c_1(g_0^2)$  and the one-loop expansion of  $u_0$  [78] because of the requirement  $c_1(g_0^2) = c(g_0^2)[c_1(g_0^2) + \mathcal{O}(g_0^2)]$ . The factor  $c(g_0^2)$  can be chosen freely, as it can be absorbed in the bare coupling constant due to the fact that in the path integral the action is divided by a factor of  $g_0^2$ .

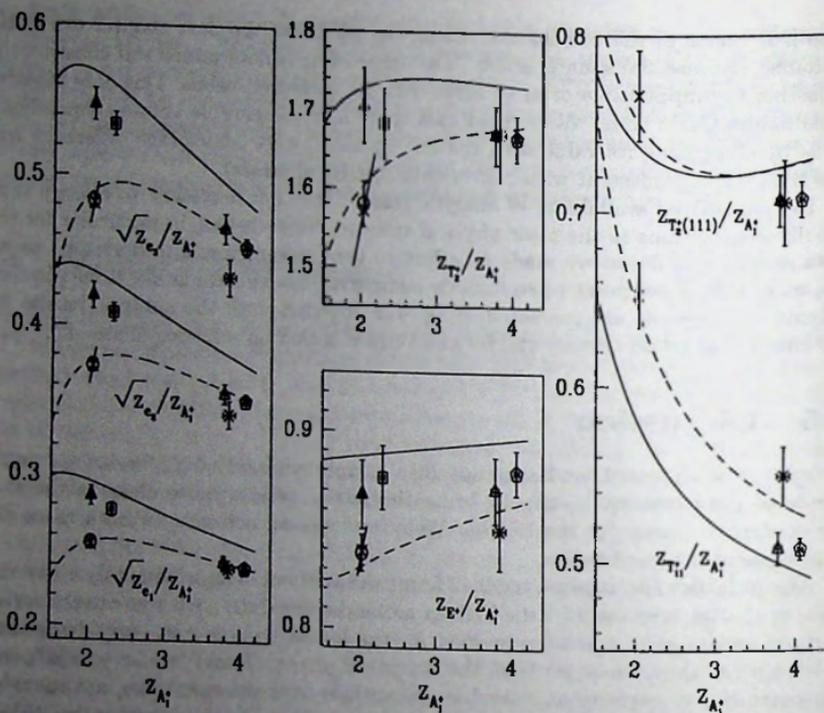


Figure 4-2. SU(2) Monte Carlo data on a  $4^3 \times 128$  lattice for the Wilson action (circles for our data and crosses for data by Michael [15], with tilted error bars when data overlap), the Lüscher-Weisz Symanzik-improved action (triangles), the square Symanzik action (squares), and the tadpole-improved one-loop Lüscher-Weisz Symanzik action (pentagons), at lattice spacings of approximately 0.02 and 0.12 fermi. A comparison is made with analytic results for the continuum (solid lines) and Wilson action on a lattice of size  $4^3 \times \infty$  (dashed lines).

One might object that the lattice spacing we have used to study tadpole improvement,  $a = 0.12$  fermi, is not really large enough. We have certainly not probed lattice spacings as large as  $a = 0.4$  fermi, that have been advertised [31]. Nevertheless, for  $a = 0.12$  fermi,  $u_0^4 = 0.6819(1)$  significantly deviates from 1. The correction to the  $2 \times 1$  Wilson loop coefficient  $c_1$  (see eq. (4.2.1)) at these parameters is 27% with respect to its tree-level value (without tadpole improvement it would have been 17%).

## 5 One-loop Symanzik coefficients

### 5.1 Introduction

When studying pure lattice gauge theory it is important to realize that the lattice spacing  $a$  is a rather elusive quantity. The action, when formulated in terms of the natural link variables  $U_\mu(x)$ , does not contain any  $a$  dependence. The reason is that the continuum action is scale invariant. In the previous chapters we introduced the lattice spacing  $a$  in a somewhat artificial way, typically by parametrizing

$$U_\mu(x) = \exp(aA_\mu(x)) \quad (5.1.1)$$

and insisting that  $A_\mu$  is the (dimensionful!) continuum vector potential. This is only natural in a classical context where  $A_\mu$  is considered an external field, like the classical instanton field in chapter 2. In a quantum mechanical context however, defined by the path integral (with coupling constant  $g_0$ )

$$Z = \int \mathcal{D}U \exp(-S_{\text{Lat}}[U]/g_0^2), \quad (5.1.2)$$

it is merely a convenient way to make contact with other regularizations of the continuum theory. The introduction of  $A_\mu$  is not at all necessary for the extraction of physical quantities, for example by means of a Monte Carlo simulation of eq. (5.1.2).

Nevertheless, the notion of a lattice spacing is sensible in the quantum mechanical context too. The reason is that eq. (5.1.2) dynamically produces a length scale, namely the correlation length  $\xi$ . More generally, it produces masses of particles,  $m_i$ . Of course due to scale invariance only dimensionless ratios appear,  $\bar{m}_i \equiv am_i$  ( $\bar{m}_i$  is referred to as the mass in lattice units). This means that the continuum limit is very clearly defined:  $\bar{m}_i \rightarrow 0$ . Moreover, the proximity to the continuum can be assessed by considering ratios  $\bar{m}_j/\bar{m}_i$  as a function of  $a$ . For small  $a$  (i.e. small  $\bar{m}_i$ ) one thus expands

$$\frac{\bar{m}_j}{\bar{m}_i} = \frac{m_j}{m_i} + \text{Rest}_{ij}, \quad \text{Rest}_{ij} = c_{ij}^{(1)} a^2 + c_{ij}^{(2)} a^4 + \dots, \quad (5.1.3)$$

where the coefficients  $c_{ij}^{(n)}$  are independent<sup>1</sup> of  $a$ . Incidentally, in a finite volume with extent  $L$  one can of course also take ratios with the mass parameter  $2\pi/L$ .

For the conventional Wilson action [7] all coefficients  $c_{ij}^{(n)}$  generically are non-zero. Improvement means reducing the lattice artefacts ('Rest') for all values of  $i$  and  $j$  that are associated with physical masses. By Wilson's renormalization group argument [5] it should be possible to find a lattice action that produces  $\text{Rest} = 0$ .

<sup>1</sup>For convenience we neglect logarithmic corrections for the moment.

In a certain approximation, the implementation of this approach has been put on a practical level by Hasenfratz and Niedermayer et al. [23]. In this chapter we follow the less ambitious perturbative approach due to Symanzik [24].

Symanzik improvement amounts to expanding 'Rest' with respect to  $a$ , and systematically removing the  $a^2, a^4, \dots$  terms. Usually only the removal of the  $a^2$  term is pursued, and we will do so in this chapter. From now on we mean by Symanzik improvement this limited reduction only. For pure gauge theories, Symanzik improvement is believed to be attainable [25, 79] by adding to the standard Wilson action a finite number of extra terms with coefficients  $c_i(g_0^2)$ , cf. eq. (4.2.1). Due to asymptotic freedom [1], the continuum limit  $a \rightarrow 0$  is achieved by approaching the critical point,  $g_0 \rightarrow 0$ . Therefore it is sensible to calculate  $c_i(g_0^2)$  perturbatively as a function of  $g_0^2$ . To this end one expands

$$c_i(g_0^2) = \sum_{m=0}^{\infty} c_i^{(m)}(g_0^2)^m. \quad (5.1.4)$$

Symanzik improvement to  $n$  loops amounts to fixing  $c_i^{(0)} \dots c_i^{(n)}$  such that physical quantities show only  $\mathcal{O}(a^4, a^2 g_0^{2(n+1)})$  deviations from their continuum values. At tree level ( $n = 0$ ) the analysis is rather easy, and many alternative Symanzik-improved actions have been introduced. However, for a long time the availability of a one-loop Symanzik-improved action has been limited to the Lüscher-Weisz action [26]. This chapter deals with the computation of the one-loop coefficients for a second case, namely the square action that was introduced in the previous chapter.

In present-day Monte Carlo simulations  $g_0^2$  is typically of order 1, and this value seems to be too large for Symanzik improvement to be effective. For this reason a phenomenological extension of Symanzik improvement, called tadpole improvement, was introduced [29]. In this approach a mean-field parameter [30]  $u_0 \approx \langle U_\mu(x) \rangle$  is included in a tree-level or one-loop Symanzik-improved action by dividing each link variable by  $u_0$ . This has been argued [29, 31] to capture most of the higher-order corrections in  $g_0$ , thus approximating an all-loop Symanzik-improved action. In practice, Monte Carlo data for tadpole-improved actions often, but not always, show dramatic improvement. As a side remark, let us mention that the goal of an *exact* all-loop Symanzik action might be attainable to sufficient precision by using Monte Carlo techniques. In ref. [28] this has been studied for the fermionic sector of QCD, where the leading lattice artefacts to be cancelled are of the order  $a$  instead of  $a^2$ .

We believe that our introduction of an alternative one-loop Symanzik-improved action is useful for several reasons. It allows one to make the following analyses of improvement:

- Estimate remaining lattice artefacts by comparing the effectiveness of two different Symanzik-improved actions.
- Study the universality and consistency of tadpole improvement.

(At tree-level such analyses were performed in, e.g., ref. [31]). Moreover, a one-loop calculation has some spin-offs:

- The Lambda parameter ratio  $\Lambda_{\text{improved}}/\Lambda_{\text{Wilson}}$ , which is needed for doing simulations with both actions at the same physical scale.
- The tadpole parameter  $u_0$  to one-loop order.

The outline of this chapter is as follows. In section 5.2 we discuss the notion of *on-shell* improvement [25]. We will see that only two improvement conditions need to be satisfied in order to cancel the  $\mathcal{O}(a^2)$  errors in all physical quantities. In section 5.3 we set up lattice perturbation theory. All propagators and vertices needed are listed in the appendices A–C. Then in section 5.4 we calculate the static quark potential to one-loop order. This allows us to extract the Lambda and tadpole parameters, and also one improvement condition. In section 5.5 we discuss perturbation theory in a twisted finite volume, which subsequently is used to obtain two improvement conditions and, again, the value of the Lambda parameter. We conclude with section 5.6, where the main results are listed in a compact way and a simple test of tadpole improvement is performed.

Our calculation follows closely the well-documented work of Lüscher, Weisz and Wohlert [25–27, 78, 79]. Therefore we only go into full detail at points where we take a different approach, and at some points where the earlier treatments can be complemented. Special attention will be paid to the many checks we have performed to convince ourselves of the correctness of the final results.

## 5.2 On-shell improvement

Up to now we have not specified precisely what quantities are to be improved. In Symanzik's work [24] the aim was to improve correlation functions. However, correlation functions depend on the elementary field operators. As a result it is difficult to find a formulation in terms of which this kind of improvement can be implemented. For the  $O(3)$   $\sigma$ -model, Symanzik succeeded by performing a well-chosen field redefinition, but for gauge theories the problem becomes more intractable and to our knowledge no one has solved it up to date.

One can avoid this complication by restricting oneself to 'on-shell improvement' as introduced by Lüscher and Weisz [25]. Instead of improving correlation functions one now pursues only the improvement of spectral quantities. Unlike correlation functions these are necessarily invariant under field redefinitions. Examples of spectral quantities are masses of stable particles and the static quark potential.

In order to determine a minimal set of improvement conditions it is sufficient [25] to perform a small- $a$  expansion of the classical action. The most general action we consider is given by eq. (4.2.1), the expansion of which we copy from eq. (4.2.2):

$$\begin{aligned}
S(\{c_i(g_0^2)\}) &= -\frac{a^4}{2}(c_0 + 8c_1 + 8c_2 + 16c_3 + 16c_4)(g_0^2) \sum_{x,\mu,\nu} \text{Tr} (F_{\mu\nu}(x))^2 \\
&+ \frac{a^6}{12}(c_0 + 20c_1 - 4c_2 + 4c_3 + 64c_4)(g_0^2) \sum_{x,\mu,\nu} \text{Tr} (D_\mu F_{\mu\nu}(x))^2 \\
&+ \frac{a^6}{3}(c_2 + 3c_3)(g_0^2) \sum_{x,\mu,\nu,\lambda} \text{Tr} (D_\mu F_{\mu\lambda}(x) D_\nu F_{\nu\lambda}(x)) \\
&+ \frac{a^6}{3} c_2 (g_0^2) \sum_{x,\mu,\nu,\lambda} \text{Tr} (D_\mu F_{\nu\lambda}(x))^2 + \mathcal{O}(a^8). \tag{5.2.1}
\end{aligned}$$

In this equation,  $F_{\mu\nu} = \partial_\mu \bar{A}_\nu - \partial_\nu \bar{A}_\mu + [\bar{A}_\mu, \bar{A}_\nu]$  and  $D_\mu = \partial_\mu + \text{ad} \bar{A}_\mu$ , where  $\bar{A}_\mu$  is related to  $U_\mu$  by a path ordered exponentiation (see eq. (C.36)). A crucial observation is that  $c_4(g_0^2)$  appears only in the combinations

$$\bar{c}_0(g_0^2) \equiv c_0(g_0^2) - 16c_4(g_0^2), \quad \bar{c}_1(g_0^2) \equiv c_1(g_0^2) + 4c_4(g_0^2). \tag{5.2.2}$$

This implies that with the substitutions  $c_0(g_0^2) \rightarrow \bar{c}_0(g_0^2)$ ,  $c_1(g_0^2) \rightarrow \bar{c}_1(g_0^2)$ , the entire analysis of ref. [25] can be copied literally. The only subtlety is that positivity of the lattice action must be rechecked for  $c_4(g_0^2) \neq 0$ . Postponing this proof to appendix E, we take over the following conclusions from ref. [25]:

- The normalization condition

$$\bar{c}_0(g_0^2) + 8\bar{c}_1(g_0^2) + 8c_2(g_0^2) + 16c_3(g_0^2) = 1 \tag{5.2.3}$$

can be imposed. This ensures that eq. (5.2.1) has a natural continuum limit and, more importantly, determines  $\bar{c}_0(g_0^2)$  as a function of the other coefficients. Since in the path integral the action is weighted by a factor of  $1/g_0^2$ , eq. (5.2.3) is always attainable by a redefinition of the bare coupling constant  $g_0$ .

- After imposing the normalization condition, on-shell improvement is expected to be satisfied for all spectral quantities once the coefficients  $\bar{c}_1(g_0^2)$ ,  $c_2(g_0^2)$  and  $c_3(g_0^2)$  are tuned correctly. Furthermore,  $c_3(g_0^2)$  can be chosen 0 because of the freedom of field redefinitions. Hence in fact only  $\bar{c}_1(g_0^2)$  and  $c_2(g_0^2)$  are to be determined, i.e. we need only two improvement conditions.
- The tree-level conditions for on-shell improvement are

$$\begin{aligned}
\bar{c}_1 &= -\frac{1}{12}, \\
c_2 &= 0. \tag{5.2.4}
\end{aligned}$$

(This will be verified along the way in this chapter).

Moreover, the fact that  $c_0(g_0^2)$ ,  $c_1(g_0^2)$  and  $c_4(g_0^2)$  enter the analysis only through the combinations  $\bar{c}_0(g_0^2)$  and  $\bar{c}_1(g_0^2)$ , implies that one of them, e.g.  $c_4(g_0^2)$ , is a free function. Having noted this, one should yet keep in mind that the one-loop coefficients<sup>2</sup>  $\bar{c}'_0$ ,  $\bar{c}'_1$ , and  $c'_2$  do depend on the tree-level choice of  $c_4$ . In particular, the Lüscher-Weisz

<sup>2</sup>We use the notation  $c_i = c_i^{(0)}$ ,  $c'_i = c_i^{(1)}$  throughout this chapter.

and square actions will have different one-loop coefficients. Nonetheless  $c_4'$  is a free parameter for both actions.

In the remainder of this chapter we limit ourselves to actions satisfying eq. (5.2.3) and  $c_3(g_0^2) = 0$ , though some intermediate results are valid more generally. In particular we focus on the Wilson action (abbreviation W;  $c_0 = 1, c_1 = c_2 = c_3 = c_4 = 0$ ) as well as the Lüscher-Weisz (LW;  $c_0 = \frac{5}{3}, c_1 = -\frac{1}{12}, c_2 = c_3 = c_4 = 0$ ) and square (sq;  $c_0 = \frac{15}{9}, c_1 = -\frac{1}{6}, c_2 = c_3 = 0, c_4 = \frac{1}{144}$ ) Symanzik actions. For the last two actions we determine the coefficients  $c_1'$  and  $c_2'$  in sections 5.4 and 5.5.

### 5.3 Lattice perturbation theory

This will be a very short description of lattice perturbation theory. For a thorough setup of the general formalism see refs. [20, 27].

As we saw in chapter 1, the elementary field in pure lattice gauge theories is the link variable  $U_\mu(x)$ . In a perturbative approach it is convenient to use the parametrization (5.1.1). Since  $U_\mu(x) \in \text{SU}(N)$ ,  $A_\mu(x)$  lies in the Lie algebra of  $\text{SU}(N)$ :

$$A_\mu(x) = g_0 \sum_a A_\mu^a(x) T^a, \quad (5.3.1)$$

where  $A_\mu^a(x) \in \mathbb{R}$ . We limit ourselves to the fundamental representation, with conventions

$$\text{Tr} (T^a T^b) = -\frac{1}{2} \delta_{ab}, \quad [T^a, T^b] = f_{abc} T^c. \quad (5.3.2)$$

Note that eq. (5.1.1) contains an ordinary exponential, not a path ordered exponential. Therefore eq. (5.2.1) receives corrections. In appendix C this provides us with a non-trivial check on the vertices.

Using eq. (5.1.1) it is straightforward to expand the action (4.2.1) to a given order in  $g_0$ . The only complication is the many terms involved, but this can be handled by a symbolic computer language such as FORM [56] or Mathematica [61].

Another issue is gauge fixing. We have adopted the gauge choice of ref. [78], i.e. covariant gauge fixing. The advantage of this choice over the Coulomb-like choice in ref. [27] is that the gauge condition,

$$\sum_\mu \partial_\mu^L A_\mu(x) \stackrel{\text{covariant g.f.}}{=} 0, \quad (5.3.3)$$

is independent of the lattice action chosen, thus eliminating a potential source of mistakes when changing from the LW action to the square action. Here  $\partial^L$  is the left lattice derivative,  $\partial_\mu^L f(x) \equiv (f(x) - f(x - a\hat{\mu}))/a$ . We implement eq. (5.3.3) by adding to the action

$$S_{\text{gf}} = -\frac{1}{\alpha} a^4 \sum_x \text{Tr} \left( \sum_\mu \partial_\mu^L A_\mu(x) \right)^2, \quad (5.3.4)$$

where  $\alpha$  is the well-known gauge parameter that should drop out of physical quantities.

Of course the gauge fixing procedure induces ghosts in the usual way. Hence a ghost term  $S_{\text{ghost}}$  must be added to the action. One more term to be added is the Haar measure term  $S_{\text{measure}}$  associated with the parametrization (5.1.1), (5.3.1). The total action thus reads

$$\frac{1}{g_0^2} S_{\text{total}} = \frac{1}{g_0^2} (S(\{c_i(g_0^2)\})) + S_{\text{gf}} + S_{\text{measure}} + S_{\text{ghost}}. \quad (5.3.5)$$

For the calculations in the following sections we need the expansion of  $S_{\text{total}}/g_0^2$  up to third order in  $g_0$ . It is given in the appendices A–C. Using this expansion, and the Euclidean path integral

$$Z = \int \exp\left(-\frac{1}{g_0^2} S_{\text{total}}\right), \quad (5.3.6)$$

perturbation theory is set up like in the continuum, with the exception that coordinates are now discrete or, equivalently, momenta run over one Brillouin zone only. Correspondingly, momentum-conserving delta functions  $\delta(k_1 + \dots + k_n)$  are  $2\pi/a$  periodic. In the remaining sections we use lattice units. The  $a$  dependence can be reobtained from a dimensional analysis.

## 5.4 Static quark potential

### 5.4.1 Generalities

This section is devoted to the calculation of the static quark potential to one-loop order.

The static quark potential,  $V(L)$ , is often used for the study of quark confinement within the framework of pure gauge theory. In its definition, a macroscopic Wilson loop acts as an external quark source (for further justification see e.g. ref. [20]):

$$V(L) = \lim_{T \rightarrow \infty} -\frac{1}{T} \ln \left( \frac{1}{N} \langle \text{Re Tr } U(L \times T) \rangle \right). \quad (5.4.1)$$

Here  $U(L \times T)$  denotes a rectangular Wilson loop with extension  $T$  in the time direction ( $\hat{0}$ ) and  $L$  in some spatial direction, say  $\hat{1}$ ; see figure 5-1. The expectation value  $\langle \dots \rangle$  denotes averaging with respect to the path integral (in infinite volume). Since  $V(L)$  captures the exponential decay of a gauge invariant correlation function, it can in principle be expressed as an ( $L$  dependent) function of the spectrum. Therefore its small- $a$  behavior, or in lattice units large- $L$  behavior, is suitable for a (partial) determination of the on-shell improvement coefficients.

To avoid confusion, note that we will not pursue the computation of  $V(L)$  in the confining regime, i.e. at large *physical* distances  $L \gg \xi$  (where  $\xi$  is the correlation length). That would not be feasible within the framework of ordinary perturbation

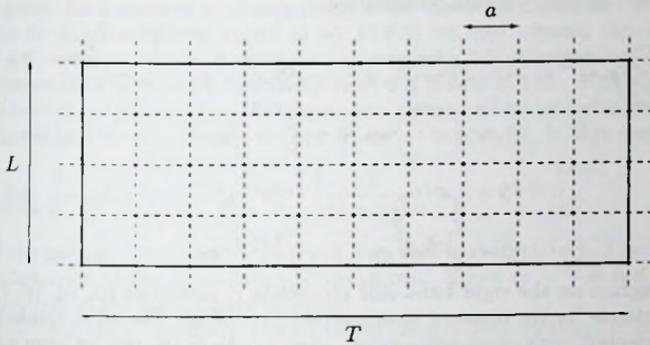


Figure 5-1.  $L \times T$  Wilson loop on the lattice, representing the path ordered product of all link variables along the loop.

theory. However, it is the basic postulate of Symanzik improvement that all physical quantities can be improved simultaneously, so that considering short-distance quantities only is sufficient for a complete determination of the improvement coefficients.

The potential  $V(L)$  not only depends on  $L$ , but also on the coupling  $g_0$ . Roughly following the notation of refs. [78, 79], but limiting ourselves to the fundamental representation of the Wilson loop, we expand

$$-\ln \left( \frac{1}{N} \text{Re Tr } U(L \times T) \right) = \sum_{n=1}^{\infty} \frac{g_0^{2n}}{(2n)!} w_n(L, T). \quad (5.4.2)$$

We restrict ourselves to the tree-level and one-loop contributions ( $n = 1, 2$ ). In terms of Feynman diagrams they read

$$w_1 = - \left[ \text{Diagram 1} \right], \quad (5.4.3)$$

$$w_2 = \left( 3w_1^2 - \left[ \text{Diagram 2} \right] \right) - 4 \left[ \text{Diagram 3} \right] - 12 \left[ \text{Diagram 4} \right]. \quad (5.4.4)$$

The endpoints of the gluon lines (  $\sim$  ) represent source terms<sup>3</sup> due to the Wilson loop, and hence are to be summed over all lattice sites on the loop. On the other

<sup>3</sup>These are proportional to  $g_0$ , cf. eqs. (5.1.1), (5.3.1).

hand, vertices ( $\bullet$ ) should be summed over all sites in space-time. The striped graph represents the one-loop vacuum polarization:

$$\begin{aligned}
 \text{Striped Loop} &= \text{Wavy line with } i \text{ in a circle} + \text{Wavy line with } x \text{ in a circle} + \text{Wavy line with dashed circle} \\
 &+ \text{Wavy line with dashed circle and wavy line} + \text{Wavy line with star-shaped loop} \\
 &+ \text{Wavy line with star-shaped loop } V + \text{Wavy line with star-shaped loop } W.
 \end{aligned}
 \tag{5.4.5}$$

The first diagram on the right hand side represents  $c_1^4$  insertions (cf. eq. (A.7)), the second one stands for the measure term insertion, eq. (A.3). The other symbols have the usual meaning, with ghost propagator  $\dashrightarrow$ . As in ref. [78] we have split the four-vertex in two parts, corresponding to the second and third lines of eq. (A.10).

The analytic expressions corresponding to the graphs in eqs. (5.4.3)–(5.4.5) can be found in ref. [78], both for finite  $T$  and for  $T \rightarrow \infty$ , and in terms of arbitrary coefficients<sup>4</sup>  $c_i$ . Below we copy those expressions whenever needed.

At tree level the following result was found [78, 79]:

$$\lim_{T \rightarrow \infty} \frac{1}{T} w_1(L, T) = 4C_f \int_{\mathbf{k}} \sin^2(\frac{1}{2}k_1 L) D_{00}(k) |_{k_0=0}, \tag{5.4.6}$$

where

$$\int_{\mathbf{k}} \equiv \prod_{j=1}^3 \left( \int_{-\pi}^{\pi} \frac{dk_j}{2\pi} \right), \tag{5.4.7}$$

$C_f$  is the Casimir in the fundamental representation,

$$C_f = \frac{1}{2} \left( N - \frac{1}{N} \right), \tag{5.4.8}$$

and  $D_{\mu\nu}(k)$  is the gauge field propagator, see eq. (A.8) and appendix B. Since  $\mu = \nu = 0$  and  $k_0 = 0$ , the form (B.14) of the propagator is applicable:

$$\begin{aligned}
 D_{00}^{-1}(k) &\stackrel{k_0=0}{\equiv} \hat{\mathbf{k}}^2 - (\bar{c}_1 - c_2) \hat{\mathbf{k}}^{(4)} - c_2 (\hat{\mathbf{k}}^2)^2 \\
 &= \mathbf{k}^2 - (\bar{c}_1 - c_2 + \frac{1}{12}) \mathbf{k}^{(4)} - c_2 (\mathbf{k}^2)^2 + \mathcal{O}(k^6).
 \end{aligned}
 \tag{5.4.9}$$

The symbol  $\hat{\mathbf{k}}^{(4)}$  and similarly  $\mathbf{k}^{(4)}$  are defined in eq. (B.5), however now summing over spatial indices only.

From eq. (5.4.9) one might be tempted to conclude that the static quark potential at tree level gives rise to two improvement conditions,  $\bar{c}_1 - c_2 = -\frac{1}{12}$  and  $c_2 = 0$ . This conclusion is however incorrect [25] (though the values are correct, cf. eq. (5.2.4))

<sup>4</sup>In refs. [78, 79],  $c_2 \leftrightarrow c_3$  with respect to the convention of refs. [25–27] and this thesis.

because  $\mathbf{k}$  is an integration variable. Since the issue comes up again at one-loop level, let us pause for a moment to clearly analyze the tree-level situation.

Due to the momentum cutoff in eq. (5.4.7) we can expand the integrand in eq. (5.4.6) with respect to  $\bar{k}_j$  without running into UV divergences. This is useful because each factor of  $k$  effectively carries a factor of  $1/L$ . Furthermore,  $V(L)$  is a potential, so we can drop  $L$ -independent terms and hence, also using that the propagator is even in  $k$ , replace  $\sin^2(\frac{1}{2}k_1L)$  by  $-\frac{1}{2}\exp(ik_1L)$ . In this way we obtain

$$\lim_{T \rightarrow \infty} \frac{1}{T} w_1(L, T) = -2C_f \int_{\mathbf{k}} e^{ik_1L} \left[ \frac{1}{\mathbf{k}^2 + \alpha \bar{\mathbf{k}}^{(4)}} + \beta + \mathcal{O}(k^2) \right], \quad (5.4.10)$$

where the parameters  $\alpha$  and  $\beta$  equal  $-\bar{c}_1 + c_2$  and  $c_2$  respectively. Due to the oscillatory behavior of the phase factor,  $\beta$  drops out. Hence  $c_2 = 0$  is *not* required for improvement of the static quark potential.

By writing the  $k_1$  integration as a contour integral over  $\exp(ik_1)$  it is not difficult to integrate the other terms analytically. The result reads [25]

$$\lim_{T \rightarrow \infty} \frac{1}{T} w_1(L, T) = -\frac{C_f}{2\pi L} \left[ 1 - 3 \left( \frac{\alpha - \frac{1}{12}}{L^2} \right) + \mathcal{O}(L^{-4}) \right]. \quad (5.4.11)$$

The conclusion is that  $\alpha = \frac{1}{12}$ , i.e.  $\bar{c}_1 - c_2 = -\frac{1}{12}$ , is the only tree-level condition for improvement of the static quark potential.

Now let us proceed to the one-loop level. In ref. [78] it is shown that

$$\lim_{T \rightarrow \infty} \frac{1}{T} w_2(L, T) = 48C_f \int_{\mathbf{k}} \sin^2(\frac{1}{2}k_1L) \bar{w}_2(\mathbf{k}) (D_{00}(\mathbf{k})|_{k_0=0})^2, \quad (5.4.12)$$

where the  $L$ -independent function  $\bar{w}_2(\mathbf{k})$  has the expansion<sup>5</sup>

$$\bar{w}_2(\mathbf{k}) = \mathbf{k}^2 \left\{ \bar{a}_1 - \beta_0 \ln \mathbf{k}^2 + (\bar{a}_2 + \bar{b}_2 \ln \mathbf{k}^2) \frac{\mathbf{k}^{(4)}}{\mathbf{k}^2} + (\bar{a}_3 + \bar{b}_3 \ln \mathbf{k}^2) \mathbf{k}^2 + \mathcal{O}(k^4) \right\}. \quad (5.4.13)$$

The coefficients  $\bar{a}_n, \bar{b}_n$  depend on the tree-level values  $\bar{c}_1$  and  $c_2$ . In particular,  $\bar{b}_2$  and  $\bar{b}_3$  are zero<sup>6</sup> if eq. (5.2.4) holds, which from now on we will assume. On the other hand,

$$\beta_0 = \frac{11N}{3(4\pi)^2} \quad (5.4.14)$$

takes precisely the universal value required for renormalization (see eq. (5.4.20) below). In order to express the improvement condition in a clean way we perform the

<sup>5</sup>Throughout this chapter we adopt the convention that  $\mathcal{O}(a^n)$  may also stand for  $a^n \ln a$  in terms. Note that in lattice units,  $a$  is absorbed in  $k$ .

<sup>6</sup>This is to be expected from Symanzik's work on scalar theories [24]. For the present situation Weisz and Wohlert [78] have shown it explicitly for  $c_4 = 0$ . Extension of their result to other tree-level improved actions is not difficult, as we will see in subsection 5.4.2.

integration in eq. (5.4.12). Like in the tree-level case,  $\bar{a}_3$  does not contribute to the  $L$  dependence of  $w_2$ , and the result reads (up to an irrelevant constant)

$$\lim_{T \rightarrow \infty} \frac{1}{T} w_2(L, T) = \frac{6C_I}{\pi L} \left[ \beta_0 \ln \left( \frac{a^2}{L^2} \right) - \bar{a}_1 + \beta_0 c - 3\bar{a}_2 \frac{a^2}{L^2} + \mathcal{O} \left( \frac{a^4}{L^4} \right) \right]. \quad (5.4.15)$$

For clarity we reinstated the  $a$  dependence. The constant

$$c \equiv 4\pi \int_{\mathbb{R}^3} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{\ln \mathbf{k}^2}{\mathbf{k}^2} e^{i\mathbf{k}_1} = -2\gamma = -1.1544313298 \dots \quad (5.4.16)$$

is independent of  $a$ ,  $L$  and  $c_1(g_0^2)$ . Note that for tree-level non-improved lattice actions eq. (5.4.15) receives additional  $\mathcal{O}(a^2/L^2)$  corrections.

The total one-loop static quark potential has a finite regularization-independent limit  $a \rightarrow 0$ ,

$$V(L) = -\frac{C_I}{4\pi L} g_R^2 \left\{ 1 + g_R^2 \left[ 3\bar{a}_2 \frac{a^2}{L^2} + \mathcal{O} \left( \frac{a^4}{L^4} \right) \right] + \mathcal{O}(g_R^4) \right\}, \quad (5.4.17)$$

in terms of the renormalized,  $L$ -dependent coupling

$$g_R^2 = g_0^2 \left\{ 1 - \left[ \beta_0 \ln \left( \frac{a^2}{L^2} \right) + \beta_0 c - \bar{a}_1 \right] g_0^2 + \mathcal{O}(g_0^4) \right\}. \quad (5.4.18)$$

Since  $V(L)$  and  $g_R$  are  $a$ -independent physical objects, we deduce from eq. (5.4.17) the one-loop improvement condition

$$\bar{a}_2 = 0. \quad (5.4.19)$$

The calculation of the  $c_i$  dependent coefficient  $\bar{a}_2$  is described in the next subsection.

Also Lambda parameter ratios can be extracted [78, 80–82]. As  $g_R$  must be independent of both  $a$  and the regularization prescription, it follows from eq. (5.4.18) that  $g_0$  depends on  $a$  and  $\bar{c}_1$ ,  $c_2$ . This is expressed by the well-known renormalization group flow

$$-\frac{1}{\beta_0 g_0^2} \stackrel{a \rightarrow 0}{=} \ln(a^2 \Lambda^2) + \mathcal{O}(\ln[-\ln(a^2 \Lambda^2)]), \quad (5.4.20)$$

where the Lambda parameter  $\Lambda$  depends on the lattice action:

$$\frac{\Lambda^*}{\Lambda} = e^{\frac{1}{2\beta_0} (\bar{a}_1(\{c_i\}) - \bar{a}_1(\{c_i^*\})}. \quad (5.4.21)$$

In this equation  $\{c_i\}$  and  $\{c_i^*\}$  represent any two choices of tree-level coefficients that satisfy the normalization condition (5.2.3), and  $\Lambda$ ,  $\Lambda^*$  are the corresponding Lambda parameters.

As a final remark, also Wilson loops for finite values of  $L$  and  $T$  are objects of interest. They allow for a gauge invariant definition of the mean field or tadpole parameter  $u_0$  [29]:

$$u_0 \equiv \left( \frac{1}{N} (\text{Re Tr } U(1 \times 1)) \right)^{\frac{1}{4}}. \quad (5.4.22)$$

	Lüscher-Weisz	square
$\beta(1, 1)$	0.366262680(2)	0.3587838551(1)
$\beta(1, 2)$	0.662626785(2)	0.6542934512(1)
$\beta(2, 2)$	1.098143594(2)	1.0887235337(1)

Table 5-1. Several numerical values of eq. (5.4.25) for the Lüscher-Weisz and square actions. Note that  $\beta(L, T) = \beta(T, L)$ .

More generally one might define for  $L, T \in \mathbb{N}$

$$u_0(L, T) \equiv \left( \frac{1}{N} \langle \text{Re Tr } U(L \times T) \rangle \right)^{\frac{1}{\pi(L+T)}}, \quad (5.4.23)$$

but the choice  $L = T = 1$  is probably optimal. The reason is that if  $L$  or  $T$  becomes larger (i.e. closer to the correlation length),  $u_0(L, T)$  may start picking up physical signals instead of the lattice artefacts that tadpole improvement aims to correct for. Nevertheless we will also compute  $w_1(1, 2)$  and  $w_1(2, 2)$  because for the LW and square actions they are related to the vacuum polarization.

### 5.4.2 Particulars

We continue with the details of the one-loop calculation. It should be kept in mind that we will only compare the Wilson, Lüscher-Weisz and square actions, and therefore we set to zero all tree-level coefficients except  $c_0$ ,  $c_1$  and  $c_4$ , and all one-loop coefficients except  $c'_0$ ,  $c'_1$ ,  $c'_2$  and  $c'_4$ .

As an appetizer let us start with finite Wilson loops. From ref. [78] we read

$$w_1(L, T) = C_f \beta(L, T), \quad (5.4.24)$$

with

$$\beta(L, T) = \int_k \left( \frac{\sin \frac{1}{2} k_1 L}{\sin \frac{1}{2} k_1} \right)^2 \left( \frac{\sin \frac{1}{2} k_0 T}{\sin \frac{1}{2} k_0} \right)^2 D_{1010}(k). \quad (5.4.25)$$

$D_{\mu\nu\lambda\rho}$  is the propagator of  $-i(\partial_\sigma A_\tau - \partial_\tau A_\sigma)$ ,

$$D_{\mu\nu\lambda\rho} = \{ [\hat{k}_\mu \hat{k}_\lambda D_{\nu\rho} - (\mu \leftrightarrow \nu)] - [\lambda \leftrightarrow \rho] \}, \quad (5.4.26)$$

and  $\int_k$  is the four-dimensional analogue of eq. (5.4.7). Upon insertion of the expressions given in appendix B, the evaluation of  $\beta(L, T)$  for finite values of  $L$  and  $T$  can be handled to a high accuracy by standard integration routines (we used NAG [83] for all numerics in this subsection). The results for  $L, T \in \{1, 2\}$  are listed in table 5-1. The entries  $(L, T) = (1, 1)$  and  $(L, T) = (1, 2)$  for the LW action agree with the results of Weisz and Wohlert. They found  $\beta(1, 1) = 0.366262$  and  $\beta(1, 2) = 0.662624$ .

The rest of this subsection is devoted to  $\bar{w}_2(\mathbf{k})$ , defined in eq. (5.4.12). Eqs. (5.4.4) and (5.4.5) carry over in

$$\bar{w}_2(\mathbf{k}) = U^{(4)}(k) + U^{(3)}(k) + \pi_{00}(k), \quad (5.4.27)$$

$$\pi = \pi' + \pi^{\text{meas}} + \pi^{\text{gh}1} + \pi^{\text{gh}2} + \pi^{V_3} + \pi^{V_4} + \pi^W, \quad (5.4.28)$$

where  $k = (0, \mathbf{k})$ . The corresponding analytic expressions are copied below from ref. [78]. Propagators and vertices can be looked up in the appendices. For shortness we use the notation  $D = D_{00}$  and adopt the summation convention.

$$U^{(4)}(k) = ND^{-1}(k) \int_{k', k''} (2\pi)^4 \delta(k + k' + k'') \left\{ \left( \frac{1}{6} - \frac{1}{8} D^{-1}(k) D(k'') \right) - \left( 1 - D^{-1}(k) D(k'') \right) \frac{\cos \frac{1}{2} k'_0}{(2k'_0)} \frac{\partial}{\partial k'_0} \right\} D(k'), \quad (5.4.29)$$

$$U^{(3)}(k) = iND^{-1}(k) \int_{k', k''} (2\pi)^4 \delta(k + k' + k'') \times \frac{\cos \frac{1}{2} k'_0}{(2k'_0)} D_{\mu 0}(k') D_{\nu 0}(k'') V_{0\mu\nu}(k, k', k''), \quad (5.4.30)$$

$$\pi'_{\mu\nu}(k) = -(\bar{k}_\lambda \delta_{\mu\nu} - \bar{k}_\mu \delta_{\lambda\nu}) q'_{\mu\lambda}(k) \bar{k}_\lambda, \quad (5.4.31)$$

$$\pi^{\text{meas}}_{\mu\nu}(k) = -\frac{1}{12} N \delta_{\mu\nu}, \quad (5.4.32)$$

$$\pi^{\text{gh}1}_{\mu\nu}(k) = -\frac{1}{24} N \delta_{\mu\nu}, \quad (5.4.33)$$

$$\pi^{\text{gh}2}_{\mu\nu}(k) = \frac{1}{4} N \int_{k', k''} (2\pi)^4 \delta(k + k' + k'') \frac{\bar{k}_\mu \bar{k}_\nu - (\widehat{k' - k''})_\mu (\widehat{k' - k''})_\nu}{\bar{k}'^2 \bar{k}''^2}, \quad (5.4.34)$$

$$\pi^{V_3}_{\mu\nu}(k) = -\frac{1}{2} N \int_{k', k''} (2\pi)^4 \delta(k + k' + k'') D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') \times V_{\mu\lambda\rho}(k, k', k'') V_{\nu\lambda'\rho'}(k, k', k''), \quad (5.4.35)$$

$$\pi^{V_4}_{\mu\nu}(k) = \frac{1}{2} N \int_{k'} D_{\lambda\rho}(k') [V_{\lambda\rho\mu\nu}(k', -k', k, -k) - V_{\lambda\mu\rho\nu}(k', k, -k', -k)], \quad (5.4.36)$$

$$\pi^W_{\mu\nu}(k) = \frac{1}{2} d (\bar{k}_\lambda \delta_{\mu\nu} - \bar{k}_\mu \delta_{\lambda\nu}) \bar{k}_\lambda \sum_i c_i \int_{k'} K_{\lambda\mu}^{(i)}(k', -k', k, -k) D_{\mu\lambda\mu\lambda}(k'). \quad (5.4.37)$$

In eq. (5.4.31)  $q'_{\mu\nu}$  denotes  $q_{\mu\nu}$ , eq. (B.2), with, following the prescription of appendix A, tree-level coefficients  $c_i$  replaced by their one-loop counterparts  $c'_i$ . Note that because eq. (5.2.3) is satisfied at tree-level (and  $c_3(\mathcal{G}_0^2) = 0$ ),

$$c'_0 = -8c'_1 - 8c'_2 - 16c'_4. \quad (5.4.38)$$

The tensors  $K_{\lambda\mu}^{(i)}$  in eq. (5.4.37) are defined in appendix C. The unusual factor

$$d \equiv \frac{1}{12} (6C_f - N) = \left( \frac{1}{4} N - \frac{1}{3} N \right) \quad (5.4.39)$$

arises from a contraction [78] of eq. (A.13).

It is worthwhile to analyze how the above expressions lead to eq. (5.4.13). In the first place, they are completely finite for non-zero  $\mathbf{k}$  because UV divergences are regularized by the lattice (integrations run from  $-\pi$  to  $\pi$ ) and potential IR divergences are regularized by the external momentum  $\mathbf{k}$ . For  $\mathbf{k} = 0$  some of the integrals are logarithmically IR divergent, and this carries over into the  $\ln k^2$  corrections<sup>7</sup> to the polynomial behavior in eq. (5.4.13) as a function of  $k$  (we come back to this point below). As mentioned in appendix A, even (odd) vertices are parity even (odd). This also holds for the two-vertex, i.e. the propagator. From this the invariance of eqs. (5.4.29)–(5.4.37) under  $k \rightarrow -k$  can be verified explicitly. Hence  $w_2(\mathbf{k})$  contains no odd powers in  $k$ . Also cubic invariance can be checked, and therefore up to order  $k^4$  only the combinations  $\mathbf{k}^2$  and  $\mathbf{k}^{(4)}$  appear. Finally, there is no constant term because  $\pi_{\mu\nu}(0) = 0$  holds [78] due to gauge invariance while the expressions for  $U^{(3,4)}$  clearly are of order  $k^2$  due to an overall factor of  $D^{-1}(k)$ .

For use below note that  $\pi_{\mu\nu}(0) = 0$  allows us to subtract from each term on the right hand side of eq. (5.4.28) its value for  $k = 0$ .

We now embark on the computation of the various contributions to  $\bar{w}_2(\mathbf{k})$ . Note that all of them are proportional to  $N$ , except for  $\pi'_{00}$  and  $\pi_{00}^W$ . We therefore follow Weisz and Wohlert and define

$$z(\mathbf{k}) = \pi'_{00}(k)|_{k_0=0}; \quad x(\mathbf{k}) = \pi_{00}^W(k)|_{k_0=0}; \quad y(\mathbf{k}) = \frac{1}{N}(\bar{w}_2(\mathbf{k}) - x(\mathbf{k}) - z(\mathbf{k})). \quad (5.4.40)$$

$x(\mathbf{k})$  and  $z(\mathbf{k})$  are easily computable,

$$x(\mathbf{k}) = d \left\{ [c_0 \ell(1, 1) + 8c_1 \ell(1, 2) + 16c_4 \ell(2, 2)] \mathbf{k}^2 - [c_1 \ell(1, 2) + 4c_4 \ell(2, 2)] \hat{\mathbf{k}}^{(4)} \right\}, \quad (5.4.41)$$

$$z(\mathbf{k}) = (\bar{c}'_1 - c'_2) \hat{\mathbf{k}}^{(4)} + c'_2 (\mathbf{k}^2)^2. \quad (5.4.42)$$

The elaborate part of the computation is posed by  $y(\mathbf{k})$ . We can save ourselves a fair amount of work by pursuing only the following two objectives:

1. The evaluation of  $\Lambda_{\text{sq(uare)}}/\Lambda_{\text{W(ilson)}}$ . Due to eq. (5.4.21) this requires computing  $\bar{a}_1^{\text{sq}} - \bar{a}_1^W$ , where  $\bar{a}_1$  is defined in eq. (5.4.13).
2. The evaluation of the improvement condition (5.4.19), i.e. of the coefficient  $\bar{a}_2$  defined in eq. (5.4.13), for the square action. It is sufficient to compute  $\bar{a}_2^{\text{sq}} - \bar{a}_2^{\text{LW}}$ , because the value of  $\bar{a}_2^{\text{LW}}$  was computed by Weisz and Wohlert [78], and later verified by Lüscher and Weisz [27].

We thus need only consider  $y^*(\mathbf{k}) - y^{**}(\mathbf{k})$ , where '\*' and '\*\*' denote different lattice actions. This implies that we can drop  $\pi_{00}^{\text{meas}}$ ,  $\pi_{00}^{\text{gh1}}$  and  $\pi_{00}^{\text{gh2}}$ . Much more importantly, it gives us the opportunity to circumvent the logarithmic terms in eq. (5.4.13) due to the following argument.

<sup>7</sup>In non-lattice units this gives rise to the familiar  $\ln a$  UV divergences.

The most straightforward way to perform the  $k \rightarrow 0$  expansion of the integrals would be to expand the integrands. Unfortunately this method is too naive for terms that contain both propagators  $D(k')$  and  $D(k'')$ , because due to momentum conservation the expansion of  $D(k'')$  gives for each factor  $k^2$  a factor  $(k'^2 + \mathcal{O}(k'^4))^{-1}$ , thus possibly introducing IR divergences. As an example consider  $\pi_{\mu\nu}^{V_3}(k)$  (eq. (5.4.35)). For  $k = 0$  its integrand behaves as  $1/k'^2$  near  $k' = 0$ , so the integral is convergent. However, if we expand the integrand with respect to  $k^2$ , we introduce at  $\mathcal{O}(k^2)$  a logarithmic IR divergence. This would be fatal were it not that due to its IR nature the divergence is completely determined by the small- $k'$  behavior of the integrand. In particular it is the same for all lattice actions satisfying eq. (5.2.3). This implies that subtracting the integrands of two such actions produces an extra factor of  $(k'^2 + \mathcal{O}(k'^4))^{+1}$ , and the  $\mathcal{O}(k^2)$  term is finite.

For other contributions to  $y(k)$  one can give similar arguments. Therefore the naive method does yield the correct  $k^2$ -coefficients in the expansion of  $y^*(k) - y^{**}(k)$ . Furthermore, if '\*' and '\*\*' both denote tree-level improved actions, then the integrands are the same up to subleading order, and the naive method also produces the correct  $k^4$ -coefficients.

As a side remark we mention that if one is interested in  $y(k)$  itself, instead of merely its  $a_i$  dependence, one has to tackle the expansion with respect to  $k$  in a more sophisticated way. For example, one might subtract a suitable integrand that has the same IR behavior, but the integral of which is computable by continuum methods. This leads to the  $\ln k^2$  factors appearing in eq. (5.4.13). From the above discussion we then conclude that the coefficient  $\beta_0$  is universal within the class of all lattice actions satisfying the condition (5.2.3), and the coefficients  $\bar{b}_{2,3}$  are universal within the subclass of tree-level improved actions. Using Weisz and Wohlert's work [78] on the Lüscher-Weisz action it then follows that the latter values are  $\bar{b}_2 = 0$  and  $\bar{b}_3 = 0$ .

Let us denote by  $N a_i$  the contribution of  $N y(k)$  to  $\bar{a}_i$ . From eq. (5.4.9) and the above discussion we conclude that for the computation of  $a_1^{sq} - a_1^{W}$  it is sufficient to consider  $(\pi_{00}^{V_3}(k) - \pi_{00}^{V_3}(0))/N$ ,  $(\pi_{00}^{V_4}(k) - \pi_{00}^{V_4}(0))/N$ , and

$$\bar{U}^{(3)}(k) \equiv i k^2 \int_{k'} \frac{\cos \frac{1}{2} k'_0}{(2k'_0)} D_{\mu 0}(k') D_{\nu 0}(k') V_{0\mu\nu}(0, k', -k'), \quad (5.4.43)$$

$$\bar{U}^{(4)}(k) \equiv k^2 \int_{k'} \left( \frac{1}{6} - \frac{\cos \frac{1}{2} k'_0}{(2k'_0)} \frac{\partial}{\partial k'_0} \right) D(k'). \quad (5.4.44)$$

Similarly,  $a_2^{sq} - a_2^{LW}$  only receives contributions from  $(\pi_{00}^{V_3}(k) - \pi_{00}^{V_3}(0))/N$  and  $(\pi_{00}^{V_4}(k) - \pi_{00}^{V_4}(0))/N$  ( $U^{(3,4)}(k)$  only produce  $k^4$  terms of the type  $(k^2)^2$ ).

The integrands of  $\bar{U}^{(3,4)}$  are already expanded with respect to  $k$ . The expansion of  $\pi_{00}^{V_i}(k)$  can be performed rather easily, because the single internal propagator is independent of the external momentum  $k$ . We used Mathematica [61] to obtain

$$\begin{aligned}
& \frac{1}{N} (\pi_{00}^{V_3}(k) - \pi_{00}^{V_3}(0)) \\
&= \frac{\bar{k}^2}{12} \int_{k'} \{ (\bar{c}_0 + 8\bar{c}_1) [(7 - \frac{1}{2}\bar{k}_1'^2) D(k') + \bar{k}'_0 \bar{k}'_1 D_{01}(k')] \\
&\quad + \bar{c}_1 [(84 - 27\bar{k}'_0{}^2 - 60\bar{k}'_1{}^2 + 10\bar{k}'_1{}^4 + \frac{1}{2}\bar{k}'_0{}^2 \bar{k}'_1{}^2) D(k') \\
&\quad + 2(18 - 7\bar{k}'_1{}^2) \bar{k}'_0 \bar{k}'_1 D_{01}(k')] \\
&\quad + c_4 [2(-60 + 23\bar{k}'_0{}^2)(3 - \bar{k}'_1{}^2) \bar{k}'_1{}^2 D(k') \\
&\quad + 8(-12 + 5\bar{k}'_0{}^2)(-3 + \bar{k}'_1{}^2) \bar{k}'_0 \bar{k}'_1 D_{01}(k')] \} \\
&- \frac{\bar{k}^{(4)}}{12} \int_{k'} \{ \bar{c}_1 [(25 - 6\bar{k}'_0{}^2 - 10\bar{k}'_1{}^2 + 2\bar{k}'_0{}^2 \bar{k}'_1{}^2 + \frac{1}{2}\bar{k}'_1{}^4) D(k') \\
&\quad + (4 - \bar{k}'_1{}^2) \bar{k}'_0 \bar{k}'_1 D_{01}(k')] \\
&\quad + c_4 [(12 - 3\bar{k}'_0{}^2 - 120\bar{k}'_1{}^2 + 38\bar{k}'_0{}^2 \bar{k}'_1{}^2 + 38\bar{k}'_1{}^4 - \frac{1}{2}\bar{k}'_0{}^2 \bar{k}'_1{}^4) D(k') \\
&\quad + 2(12 - 5\bar{k}'_0{}^2)(4 - \bar{k}'_1{}^2) \bar{k}'_0 \bar{k}'_1 D_{01}(k')] \}. \tag{5.4.45}
\end{aligned}$$

For  $c_4 = 0$  this agrees with ref. [78].

The last ingredient needed is the expansion of  $(\pi_{00}^{V_3}(k) - \pi_{00}^{V_3}(0))/N$ . In principle the analytic expansion of the integrand is straightforward, but to our taste too tedious due to the  $k$ -dependence of the propagator(s). We therefore decided to evaluate  $\Delta\pi_{00}^{V_3,i}(k) \equiv (\pi_{00}^{V_3,i,\text{sq}}(k) - \pi_{00}^{V_3,i,\text{sq}}(0) - \pi_{00}^{V_3,i}(k) + \pi_{00}^{V_3,i}(0))/N$  ( $i = W, LW$ ) numerically for a number of values of  $\mathbf{k}$ . As explained above the result should be of the form

$$\Delta\pi_{00}^{V_3,i}(k) = \Delta a_1^{V_3,i} \mathbf{k}^2 + \Delta a_2^{V_3,i} \mathbf{k}^{(4)} + \Delta a_3^{V_3,i} (\mathbf{k}^2)^2 + \mathcal{O}(k^6). \tag{5.4.46}$$

(For  $i = W$  there are also logarithmic terms at  $\mathcal{O}(k^4)$ ). We determined  $\Delta a_1^{V_3,W}$  by choosing  $\mathbf{k} = (\varepsilon, 0, 0)$  and fitting to  $\varepsilon^2$ . For the determination of  $\Delta a_2^{V_3,LW}$  we subtracted results for  $\mathbf{k} = (\varepsilon, 0, 0)$  and  $\mathbf{k} = (\varepsilon/\sqrt{2}, \varepsilon/\sqrt{2}, 0)$ , thus eliminating the  $\mathbf{k}^2$  and  $(\mathbf{k}^2)^2$  terms. The results were fitted to  $\frac{1}{2}\varepsilon^4$ . These methods introduce numerical instabilities of the order of  $\delta/\varepsilon^2$  and  $\delta/\varepsilon^4$  respectively, where  $\delta$  is the computer precision ( $\delta \approx 10^{-14}$ ). One would like to choose  $\varepsilon$  not too big, but on the other hand subleading terms (which are of relative order  $\varepsilon^2$ ) must be kept sufficiently small. By choosing  $\varepsilon$  in the ranges 0.0001 to 0.001 and 0.01 to 0.1 respectively, we were able to extract  $\Delta a_1^{V_3}$  and  $\Delta a_2^{V_3}$  with roughly the same precision as we expect we could have attained by an analytic expansion and subsequent numerical integration.

Our final result in the  $y(\mathbf{k})$  sector is

$$\begin{aligned}
a_1^{\text{sq}} - a_1^{\text{W}} &= -0.031810197(2), \\
a_2^{\text{sq}} - a_2^{\text{LW}} &= 0.000087063(4). \tag{5.4.47}
\end{aligned}$$

As a check of our analysis and programs we also computed

$$a_1^{\text{LW}} - a_1^{\text{W}} = -0.031361443(2), \tag{5.4.48}$$

which agrees with the value found by Weisz and Wohlert [78],  $-0.03136145(1)$ . When we include the results for  $x(\mathbf{k})$  and  $z(\mathbf{k})$ , using  $\bar{\mathbf{k}}^2 = \mathbf{k}^2 - \frac{1}{12}\mathbf{k}^4 + \mathcal{O}(k^6)$  and table 5-1 (together with  $l(1,1)_{\text{Wilson}} = \frac{1}{2}$ ), we obtain for the coefficients of  $\bar{w}_2(\mathbf{k})$

$$\bar{a}_1^i - \bar{a}_1^{\text{W}} = \begin{cases} -\left(0.086580342(3)N - 0.082828348(1)\frac{1}{N}\right) & \text{for } i = \text{LW} \\ -\left(0.085608020(2)N - 0.08069673318(6)\frac{1}{N}\right) & \text{for } i = \text{sq}, \end{cases} \quad (5.4.49)$$

$$\begin{aligned} \bar{a}_2^{\text{sq}} - \bar{a}_2^{\text{LW}} &= [\bar{c}'_1 - c'_1]_{\text{sq}} - [\bar{c}'_1 - c'_1]_{\text{LW}} \\ &\quad - \left(0.002158351(4)N - 0.0033681212(1)\frac{1}{N}\right). \end{aligned} \quad (5.4.50)$$

It follows from eq. (5.4.21) that

$$\begin{aligned} \frac{\Lambda_{\text{LW}}}{\Lambda_{\text{W}}} &= \begin{cases} 4.1308935(3) & \text{for } N = 2 \\ 5.2921038(3) & \text{for } N = 3, \end{cases} \\ \frac{\Lambda_{\text{sq}}}{\Lambda_{\text{W}}} &= \begin{cases} 4.0919901(2) & \text{for } N = 2 \\ 5.2089503(2) & \text{for } N = 3. \end{cases} \end{aligned} \quad (5.4.51)$$

The values for the LW action agree with those in ref. [78]:  $4.13089(1)$  for  $N = 2$  and  $5.29210(1)$  for  $N = 3$ . Moreover, the Lambda ratios for the square action completely agree with the results in ref. [84]. These were obtained in a totally different way, namely by extending the background field method used in the previous chapter.

From eqs. (5.4.19), (5.4.50) we read off a one-loop improvement condition. We can substitute the result of Lüscher and Weisz [27] for the LW action,

$$[\bar{c}'_1 - c'_1]_{\text{LW}} = \begin{cases} -0.01100879(1) & \text{for } N = 2 \\ -0.02080086(2) & \text{for } N = 3, \end{cases} \quad (5.4.52)$$

which is consistent with the value extracted by Weisz and Wohlert [78] from the static quark potential, but more accurate. In this way we obtain

$$[\bar{c}'_1 - c'_1]_{\text{sq}} = \begin{cases} -0.00837615(2) & \text{for } N = 2 \\ -0.01544851(3) & \text{for } N = 3. \end{cases} \quad (5.4.53)$$

We conclude with the observation that the last diagram in eq. (5.4.5) gives quantitatively far the most important contribution to these values:

$$\begin{aligned} \frac{-\bar{a}_2^{\text{LW}}|_{\text{LW}}}{[\bar{c}'_1 - c'_1]_{\text{LW}}} &= \begin{cases} 0.78 & \text{for } N = 2 \\ 0.82 & \text{for } N = 3, \end{cases} \\ \frac{-\bar{a}_2^{\text{sq}}|_{\text{sq}}}{[\bar{c}'_1 - c'_1]_{\text{sq}}} &= \begin{cases} 0.69 & \text{for } N = 2 \\ 0.75 & \text{for } N = 3. \end{cases} \end{aligned} \quad (5.4.54)$$

The other diagrams give much smaller contributions, so that the large ratios are not due to coincidental cancellations. This fits well in the tadpole/mean field picture because the  $W$ -vertex is not present in the continuum theory, and hence is at least partly responsible for the deviation from 1 of the tadpole parameter  $u_0$ .

## 5.5 Spectroscopy in a twisted finite volume

### 5.5.1 Introduction and formalism

In this subsection we give a rather short overview of the formalism and use of pure gauge theory in a partly twisted space-time. For details we refer to refs. [26, 27], and also to chapter 1.

Twist [10] was introduced by 't Hooft, who observed that physical periodic boundary conditions are still satisfied if the  $A_\mu$ -field is 'twisted'. A special case is twist over a distance  $L$  in the 1 and 2 directions (with twist quantum 1):

$$A_\mu(x + L\hat{\nu}) = \Omega_\nu A_\mu(x) \Omega_\nu^{-1}, \quad (\nu = 1, 2), \quad (5.5.1)$$

where  $\Omega_\nu \in \text{SU}(N)$  are  $A_\mu$  and  $x$  independent matrices satisfying

$$\Omega_1 \Omega_2 = z \Omega_2 \Omega_1, \quad z \equiv e^{2\pi i/N}. \quad (5.5.2)$$

Choosing infinite volume in the 0 and 3 directions, this space-time was called the 'twisted tube' by Lüscher and Weisz.

Due to its lack of periodicity in the  $x_\perp \equiv (x_1, x_2)$  directions, the twisted field  $A_\mu(x)$  cannot be Fourier transformed in the usual way. However, a generalized Fourier decomposition may be introduced:

$$A_\mu(x) = g_0 \oint_k e^{ik(x + \frac{1}{2}a\mu)} \tilde{A}_\mu(k) \Gamma_k. \quad (5.5.3)$$

Note that  $\tilde{A}_\mu(k) \in \mathbb{C}$  does not possess a color index. Color degrees of freedom are absorbed into the momenta, as can be seen from the formulas below. In eq. (5.5.3) the integration symbol '\$\oint\$' stands for

$$\oint_k \equiv \frac{1}{L^2 N} \sum_{k_1} \int_{-\pi}^{\pi} \frac{dk_0}{2\pi} \int_{-\pi}^{\pi} \frac{dk_3}{2\pi}, \quad (5.5.4)$$

and  $\Gamma_k \in \text{SU}(N)$  satisfies

$$\Omega_\nu \Gamma_k \Omega_\nu^{-1} = e^{ik_\nu L} \Gamma_k, \quad (\nu = 1, 2). \quad (5.5.5)$$

The momentum components  $k_\perp \equiv (k_1, k_2)$  in eq. (5.5.4) are discretized:

$$k_\nu = mn_\nu, \quad n_\nu \in \mathbb{Z}, \quad m \equiv \frac{2\pi}{NL}, \quad (5.5.6)$$

and  $\sum_{k_\perp}$  runs over, say,  $n_\nu = 0, 1, \dots, NL - 1$ . The solution to eq. (5.5.5), which is unique up to a phase, was given in section 1.3:

$$\Gamma_k = \Omega_1^{-n_2} \Omega_2^{n_1} z^{\frac{1}{2}(n_1+n_2)(n_1+n_2-1)}, \quad (k_\nu = mn_\nu, \nu = 1, 2). \quad (5.5.7)$$

In terms of

$$\chi_k \equiv \begin{cases} 0 & \text{if } n_\nu = 0 \pmod{N}, \quad (\nu = 1, 2) \\ 1 & \text{otherwise,} \end{cases} \quad (5.5.8)$$

$$\begin{aligned} (k, k') &\equiv n_1 n'_1 + n_2 n'_2 + (n_1 + n_2)(n'_1 + n'_2), \\ \langle k, k' \rangle &\equiv n_1 n'_2 - n'_1 n_2, \end{aligned} \quad (5.5.9)$$

the matrices  $\Gamma_k$  satisfy [27]

$$\begin{cases} \Gamma_{k'} = \Gamma_k & \text{if } \chi_{k'-k} = 0 \\ \Gamma_k = \mathbf{1} & \text{if } \chi_k = 0 \\ \text{Tr } \Gamma_k = 0 & \text{if } \chi_k = 1 \\ \Gamma_k^\dagger = z^{-\frac{1}{2}\langle k, k \rangle} \Gamma_{-k} \\ \Gamma_k \Gamma_{k'} = \Gamma_{k+k'} z^{\frac{1}{2}\langle k, k' \rangle - \frac{1}{2}\langle k, k' \rangle}. \end{cases} \quad (5.5.10)$$

From these properties, eq. (5.5.3), and the fact that  $A_\mu(x)$  is in the Lie algebra of  $SU(N)$  it follows that

$$\bar{A}_\mu(k)^* = -z^{\frac{1}{2}\langle k, k \rangle} \bar{A}_\mu(-k), \quad (5.5.11)$$

$$\bar{A}_\mu(k) = 0 \quad \text{if } \chi_k = 0. \quad (5.5.12)$$

One can check from these constraints and eq. (5.5.6) that, for given  $k_0$ ,  $k_3$  and  $\mu$ ,  $\{\bar{A}_\mu(k)\}$  contains  $L^2(N^2 - 1)$  real degrees of freedom, as it should.

For later use it is convenient to define a delta function associated with eq. (5.5.4):

$$\delta(k) \equiv NL^2 \delta_{k,0}^\perp 2\pi \delta^1(k_0) 2\pi \delta^1(k_3), \quad (5.5.13)$$

where  $\delta^1$  is the ordinary 1-dimensional ( $2\pi$ -periodic) delta function, and  $\delta_{k,k'}^\perp$  is the  $NL$ -periodic 2-dimensional Kronecker delta,

$$\delta_{k,k'}^\perp \equiv \begin{cases} 1 & \text{if } n_\nu - n'_\nu = 0 \pmod{NL}, \quad (\nu = 1, 2) \\ 0 & \text{otherwise.} \end{cases} \quad (5.5.14)$$

Eq. (5.5.12) is an extremely nice property of the twisted tube, at least in perturbation theory. It implies that  $\bar{A}_\mu(0)$  is not a degree of freedom of the twisted gauge field, and therefore is not to be summed over in loop 'integrals'. In this way the twisted tube escapes any IR problems that lure in infinite or finite periodic volumes.

A strongly related effect can be seen very clearly at tree level. As follows from the Feynman rules listed in appendix D, the bare propagator equals

$$\langle \bar{A}_\mu(k) \bar{A}_\nu(k') \rangle_{g_0=0} = \delta(k+k') \left( -\frac{1}{2} z^{-\frac{1}{2}\langle k, k \rangle} \right) \chi_k D_{\mu\nu}(k). \quad (5.5.15)$$

(We slightly disagree with ref. [27]). Its mass-shell equation is determined by the ordinary Feynman propagator  $D_{\mu\nu}(k)$  and thus reads

$$k_0 = \pm i E_0(\mathbf{k}), \quad E_0(\mathbf{k}) = \sqrt{k_1^2 + k_3^2} (1 + \mathcal{O}(k^2)). \quad (5.5.16)$$

However, since  $k_{\perp}$  is discretized and  $k_{\perp} = 0$  is not permissible, a mass gap emerges, of width  $m$  (in the continuum limit<sup>8</sup>). As long as  $L$  is chosen so small (in physical units) that asymptotic freedom ensures  $g_R^2(L) \ll 1$ , the tree-level value of the mass gap receives only small quantum mechanical corrections, which are computable in perturbation theory.

In refs [26, 27], Lüscher and Weisz present a Kaluza-Klein picture of the twisted tube by viewing it as a two-dimensional theory; the compact dimensions are considered internal space. In this interpretation  $n_1$  and  $n_2$  are quantum numbers. The particles defined by  $n_{\perp} = (1, 0)$  and  $n_{\perp} = (1, 1)$  (or cubic transformations thereof) are called 'A' and 'B' mesons. Their masses are  $m_A = m$  and  $m_B = \sqrt{2}m$ , to lowest order in the lattice spacing and coupling constant. The mesons possess a 'spin' quantum number taking values  $\pm$ , corresponding to the two physical polarizations of the underlying four-dimensional theory of massless gauge bosons. Of course there are also particles with approximate masses  $2m$ ,  $\sqrt{5}m$  etc. For  $g_0 \neq 0$ , these particles, except perhaps the ones with masses approximately equal to  $2m$ , should be expected to decay into A and B mesons. The A and B mesons themselves are too light to decay (for sufficiently small coupling), and hence stable. For an understanding of the particles in terms of the electric fluxes introduced by 't Hooft, see section 1.3.

The twisted tube brings about many well-defined spectral quantities that can be used to extract the Symanzik coefficients. The simplest quantities are the A and B masses. In subsection 5.5.2 we compute to one loop the mass of the A meson with positive spin. We will see that this gives the same improvement condition as was found from the static quark potential. The reason is that in both cases the coefficients  $c_i^+$  enter via a propagator insertion. We nevertheless decided to undertake the calculation in order to become familiar with perturbation theory on the twisted tube, and also to check eq. (5.4.53).

The Symanzik coefficients  $c_1^+$  and  $c_2^+$  can be obtained separately by also computing the one-loop correction to the three-gluon vertex. For suitable external momenta this is a spectral quantity, because it is related to the elastic scattering amplitude of two A mesons [26] (which in turn is related to the energy eigenvalue of a two-particle state). The calculation is described in subsection 5.5.3.

### 5.5.2 Mass of the A<sup>+</sup> meson

The energy spectrum of (improved) lattice field theories is best defined through the transfer matrix [75]. The spectrum dictates the exponential decay of the two-point function  $\langle A_{\mu}(x)A_{\mu}(y) \rangle$  as a function of the time separation  $x_0 - y_0$ . Equivalently, the spectrum can be read off from the pole structure of  $D_{\mu\nu}^{\text{full}}(k)$ , which for the twisted tube we define as follows (cf. eq. (5.5.15)):

$$\langle \bar{A}_{\mu}(k)\bar{A}_{\nu}(k') \rangle = \delta(k+k') \left(-\frac{1}{2}z^{-\frac{1}{2}(k,k)}\right) \chi_k D_{\mu\nu}^{\text{full}}(k). \quad (5.5.17)$$

<sup>8</sup>Exact lattice formulas can be found in appendix B.

We concentrate on the A meson with positive spin. From the discussion in section 1.3 we know that to all orders in  $g_0$  it can be represented by the polarization  $\varepsilon_\mu = \delta_{\mu,1}$  at a momentum  $k = (k_0, \mathbf{k})$ ;  $\mathbf{k} = (0, m, k_3)$ . Let us denote the corresponding eigenvalue of  $D_{\mu\nu}^{\text{full}}(k)$  by  $d_{\text{full}}(k)$  ( $= D_{11}^{\text{full}}(k)$ ). The energy  $E(\mathbf{k})$  and the associated wave function renormalization  $Z(\mathbf{k})$  belonging to this  $A^+$  meson are defined through

$$d_{\text{full}}(k) = \frac{Z(\mathbf{k})}{k_0^2 + E^2(\mathbf{k})} + (\text{regular in } k_0), \quad (5.5.18)$$

where the expansion is valid for  $k_0$  close to  $E(\mathbf{k})$ . We will assume that  $E(\mathbf{k})$  is physical in the sense that  $E(\mathbf{k}) \sim a^0$  instead of  $\sim a^{-1}$ , in non-lattice units. This can be verified from our results below. It implies [75]  $E(\mathbf{k}) \in \mathbb{R}$ . For future use we note

$$Z^{-1}(\mathbf{k}) = \frac{1}{2k_0} \frac{\partial}{\partial k_0} d_{\text{full}}^{-1}(k) \Big|_{k_0 = \pm iE(\mathbf{k})} \quad (5.5.19)$$

In terms of the vacuum polarization  $\pi_{\mu\nu}$ , defined through

$$D_{\mu\nu}^{\text{full}}(k) = D_{\mu\nu}(k) + g_0^2 D_{\mu\rho}(k) \pi_{\rho\sigma}(k) D_{\sigma\nu}(k) + \mathcal{O}(g_0^2), \quad (5.5.20)$$

we find

$$d_{\text{full}}^{-1}(k) = d_{(0)}^{-1}(k) + g_0^2 d_{(1)}^{-1}(k) + \mathcal{O}(g_0^4), \quad d_{(1)}^{-1}(k) = -\pi_{11}(k). \quad (5.5.21)$$

For the one-loop energy and wave-function renormalization, defined through

$$\begin{aligned} E(\mathbf{k}) &= E_0(\mathbf{k}) + g_0^2 E_1(\mathbf{k}) + \mathcal{O}(g_0^4), \\ Z(\mathbf{k}) &= Z_0(\mathbf{k}) + g_0^2 Z_1(\mathbf{k}) + \mathcal{O}(g_0^4), \end{aligned} \quad (5.5.22)$$

the following formulas are valid:

$$E_1(\mathbf{k}) = \frac{1}{2E_0(\mathbf{k})} Z_0(\mathbf{k}) d_{(1)}^{-1}(k) \Big|_{k_0 = iE_0(\mathbf{k})}, \quad (5.5.23)$$

$$Z_1(\mathbf{k}) = -Z_0^2(\mathbf{k}) \left\{ \frac{1}{2k_0} \frac{\partial}{\partial k_0} d_{(1)}^{-1}(k) - 2E_0(\mathbf{k}) E_1(\mathbf{k}) \left( \frac{1}{2k_0} \frac{\partial}{\partial k_0} \right)^2 d_{(0)}^{-1}(k) \right\} \Big|_{k_0 = iE_0(\mathbf{k})} \quad (5.5.24)$$

For the mass calculation we put  $k_3 = 0$ , since  $m_{A^+} \equiv E(\mathbf{k} = (0, m, 0))$ . At tree level, eq. (B.14) and hence eq. (B.19) are applicable, so that

$$m_{A^+}^{(0)} = m - (\tilde{c}_1 - c_2 + \frac{1}{11})m^2 + \mathcal{O}(m^4). \quad (5.5.25)$$

Note that this is improved for the choice (5.2.4) of coefficients, as it should be. Below we assume eq. (5.2.4) is satisfied.

At one-loop level we write

$$m_{A^+} = m_{A^+}^{(0)} + g_0^2 m_{A^+}^{(1)} + \mathcal{O}(g_0^4), \quad (5.5.26)$$

so that from the above analysis it follows that

$$m_{A^+}^{(1)} = -Z_0(\mathbf{k}) \frac{\pi_{11}(\mathbf{k})}{2m_{A^+}^{(0)}} \Big|_{\mathbf{k}=(im_{A^+}^{(0)}, 0, m, 0)} \quad (5.5.27)$$

Note that this quantity only depends on  $m$  and (implicitly)  $N$ , or equivalently  $L$  and  $N$ . From eq. (B.19) we find

$$Z_0(\mathbf{k})|_{\mathbf{k}=(0, m, 0)} = 1 + \mathcal{O}(m^4). \quad (5.5.28)$$

For the extraction of the one-loop coefficients we can neglect the  $\mathcal{O}(m^4)$  correction term.

The vacuum polarization  $\pi_{\mu\nu}(\mathbf{k})$  has of course the same expansion as in section 5.4, see eqs. (5.4.5) and (5.4.28), but the explicit forms (5.4.31)–(5.4.37) change a little on the twisted tube (as follows from appendix D):

$$\pi'_{\mu\nu}(k) = -(\hat{k}_\lambda \delta_{\mu\nu} - \hat{k}_\mu \delta_{\lambda\nu}) q'_{\mu\lambda}(k) \hat{k}_\lambda, \quad (5.5.29)$$

$$\pi_{\mu\nu}^{\text{meas}}(k) = -\frac{1}{12} N \delta_{\mu\nu}, \quad (5.5.30)$$

$$\pi_{\mu\nu}^{\text{gh}1}(k) = \frac{1}{12} \delta_{\mu\nu} \oint_{k'} g_-^2(k, k') \frac{\hat{k}'^2}{k'^2}, \quad (5.5.31)$$

$$\pi_{\mu\nu}^{\text{gh}2}(k) = -\frac{1}{8} \oint_{k', k''} \delta(k + k' + k'') g_-^2(k, k') \frac{\hat{k}_\mu \hat{k}_\nu - (\widehat{k' - k''})_\mu (\widehat{k' - k''})_\nu}{\hat{k}'^2 \hat{k}''^2}, \quad (5.5.32)$$

$$\begin{aligned} \pi_{\mu\nu}^{V_2}(k) &= \frac{1}{4} \oint_{k', k''} \delta(k + k' + k'') g_-^2(k, k') D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') \\ &\quad \times V_{\mu\lambda\rho}(k, k', k'') V_{\nu\lambda'\rho'}(k, k', k''), \end{aligned} \quad (5.5.33)$$

$$\begin{aligned} \pi_{\mu\nu}^{V_4}(k) &= -\frac{1}{8} \oint_{k'} g_-^2(k, k') D_{\lambda\rho}(k') \\ &\quad \times [V_{\lambda\rho\mu\nu}(k', -k', k, -k) - V_{\lambda\mu\rho\nu}(k', k, -k', -k)], \end{aligned} \quad (5.5.34)$$

$$\begin{aligned} \pi_{\mu\nu}^W(k) &= \frac{1}{8} (\hat{k}_\lambda \delta_{\mu\nu} - \hat{k}_\mu \delta_{\lambda\nu}) \hat{k}_\lambda \sum_i c_i \oint_{k'} (\chi_{k'} + \frac{1}{2} g_-^2(k, k')) \\ &\quad \times K_{\lambda\mu}^{(i)}(k', -k', k, -k) D_{\mu\lambda\mu\lambda}(k'). \end{aligned} \quad (5.5.35)$$

In these equations

$$g_-(k, k') \equiv 2i \sin\left(\frac{\pi}{N} \langle k, k' \rangle\right), \quad (5.5.36)$$

which equals 0 if  $\chi_{k'} = 0$  or  $\chi_{k+k'} = 0$ . A reasonable check of the above formulas is that they equal their infinite-volume counterparts in leading order for  $m \rightarrow 0$  (i.e.  $L/a \rightarrow \infty$  in non-lattice units), which is what one would expect physically. The technical reason is that in this limit the factor  $g_-^2(k, k')$  is oscillating infinitely faster than the rest of the integrands, so that it can be replaced by its average value:  $-2$ . Due to the smoothness of the resulting integrand,  $\oint_{k'}$  can be replaced by  $N \int_{k'}$ , and  $\oint_{k' \chi_{k'}}$  by  $(N - N^{-1}) \int_{k'}$ .

In ref. [27] it is shown that near the continuum limit  $m_{\Lambda^+}^{(1)}$  is of the form

$$\frac{m_{\Lambda^+}^{(1)}}{m} = \frac{\bar{m}_{\Lambda^+}^{(1)}}{m} - (\bar{c}'_1 - \bar{c}'_2) m^2 + \mathcal{O}(m^4); \quad \frac{\bar{m}_{\Lambda^+}^{(1)}}{m} = a_0 + a_1 m^2 + \mathcal{O}(m^4), \quad (5.5.37)$$

where we separated the contribution of  $\pi'_{11}$  (containing all dependence on  $\bar{c}'_i$ ). The coefficients<sup>9</sup>  $a_i$  are determined by the other contributions. It is clear that the improvement condition reads

$$\bar{c}'_1 - \bar{c}'_2 = a_1. \quad (5.5.38)$$

In eq. (5.5.37) there are no quadratically ( $1/m^2$ ) or logarithmically ( $m^0 \ln m$ ) divergent terms because  $m_{\Lambda^+}^{(0)}$  is independent of  $g_0$ , while multiplicative renormalization of  $g_0$  alone is sufficient to cancel all divergences in the continuum limit. Also the  $m^2 \ln m$  term is absent due to tree-level improvement, cf. the discussion below eq. (5.4.13). Note however that individual diagrams can give  $1/m^2$ ,  $\ln m$  or  $m^2 \ln m$  contributions to  $\bar{m}_{\Lambda^+}^{(1)}/m$ . For example,  $\pi^{\text{mass}}$  contributes  $N/(24m^2)$ . In the Coulomb gauge, odd powers of  $m$  can even appear [27], but not so in the covariant gauge. The reason is that the Coulomb propagator contains  $1/k^2$  divergences for  $\mathbf{k} \rightarrow 0$ , while the covariant propagator only has  $1/k^2$  poles. In any case,  $\bar{m}_{\Lambda^+}^{(1)}/m$  being of the form (5.5.37) is a very good global check against computational errors.

Following ref. [27], we decided *not* to perform the small- $m$  expansion of eqs. (5.5.31)–(5.5.35) analytically. Instead, we computed the sum of Feynman diagrams numerically for a number of values of  $L$ , and fitted the results to the expected form (5.5.37). This is much alike to what we did in the previous section to compute the coefficients appearing in eq. (5.4.46). However, a difference of practical importance is that the CPU time needed for the evaluation of  $\bar{m}_{\Lambda^+}^{(1)}$  increases for decreasing  $m$  (see below), while for  $\Delta\pi_{00}^{V_2}(k)$  it is independent of the fit variable  $k$ . Another difference is that  $\bar{m}_{\Lambda^+}^{(1)}$  depends on  $N$ . Due to the twist it is difficult to unravel this  $N$  dependence analytically, and like in ref. [27] we did separate computations for  $N = 2$  and  $N = 3$ .

We did not use entirely the same approach as in ref. [27]. The most important differences are:

1. We used covariant instead of Coulomb gauge fixing.
2. Lüscher and Weisz fully automatized the generation of vertex subprograms. That is, they wrote meta-programs that take a Wilson loop as input, and give vertex subprograms as output. We wrote the subprograms by hand, using the vertices that we had already used before for the computation of the static quark potential. As mentioned in section 5.3, the generation of the vertices themselves was automatized by means of Mathematica [61]. As a precaution against programming errors, we always performed a number of numerical tests of our subprograms against the corresponding Mathematica representations.

A further technical difference is that Lüscher and Weisz used PL/I [85] for the generation of vertices.

<sup>9</sup>In spite of the same notation, these are different from the coefficients in subsection 5.4.2.

3. Lüscher and Weisz followed a rather unconventional way to extract the coefficients  $a_0$  and  $a_1$ . For example, to find  $a_0$  they constructed from the data for, say,  $L = 6, 8, \dots, 30$  a new data series for  $L = 8, 10, \dots, 28$  that is improved in the sense that the  $m^2$  term in eq. (5.5.37) is cancelled, leaving only  $\mathcal{O}(m^4)$  deviations. This procedure was then iterated. We preferred doing a least-squares fit, which for our data appeared to give somewhat more stable results.

Once the vertex (and propagator) subprograms are ready, the programming of eqs. (5.5.31)–(5.5.35) is easy. The main point of interest is the integration routine, since it illustrates once more the merit of the twisted tube. One should appreciate that the integration routine is required to be extraordinary accurate, because many digits are lost when extracting the coefficients, in particular  $a_1$ . We required a relative accuracy of  $10^{-13}$  or better. Whether or not an integration routine is capable of attaining such a high accuracy in a reasonable amount of time depends on the the number of integration variables and the smoothness of the integrand. Now for the twisted tube the 'integration' symbol  $\int$ , eq. (5.5.4), involves only two integration variables,  $k_0$  and  $k_3$ . The components  $k_1$  and  $k_2$  are to be summed over, costing a rather cheap<sup>10</sup> factor of  $(N^2 - 1)L^2$ . More importantly, the integrand is periodic (over one Brillouin zone) and analytic on the domain of integration of  $k_0$  and  $k_3$  (due to the mass gap, poles are shifted into the complex plane). Such a situation is ideal for constructing efficient integration routines, especially if one performs a change of variables [27] that shifts the pole further away from the domain of integration. As a result, we typically needed only 50<sup>2</sup> points to approximate  $\int_{k_1, k_2}$  within the required accuracy (Lüscher and Weisz quote 32 rather than 50, apparently due to the different pole structure in the Coulomb gauge, or a more efficient change of variables).

Concerning the analyticity of the integrand in  $k_0$  and  $k_3$ , we stress that the implementation of momentum conservation for diagrams with two internal propagators is to be chosen wisely. The reason is as follows. Due to tree-level on-shellness, the external component  $k_0$  of the momentum  $k = (k_0, 0, m, 0)$  is imaginary, and approximately equal to  $im$ . Hence if one took  $k'_0, k'_3 \in \mathbb{R}$  as the integration variables, the  $k''$  propagator would be singular for  $\{k'_0 \approx k'_3 \approx 0, k'_1 = k'_2 = -m\}$  since in that case  $k'' \approx (-im, m, 0, 0)$  and  $k''^2 \approx 0$ . It is better to shift  $k'_0 = \bar{k}_0 - \frac{1}{2}k_0$  (hence  $k'_3 = -\bar{k}_0 - \frac{1}{2}k_0$ ), where  $\bar{k}_0$  (together with  $k'_3$ ) is the  $\mathbb{R}$ -valued integration variable.

Our results for individual lattice sizes are summarized in table 5-2. Note that the cancellation of quadratic divergences decreases the accuracy by a factor of  $1/m^2$  (i.e.  $1/(am)^2$  in non-lattice units).

As mentioned above, we used least-squares fits to determine the coefficients  $a_0$  and  $a_1$ . We checked that the data is consistent with the absence of odd terms in  $m$ , and also with the absence of terms  $\ln m$  and  $m^2 \ln m$ . We then fitted the data to the form  $a_0 + a_1 m^2 + a_2 m^4 + b_2 m^4 \ln m + a_3 m^6 + b_3 m^6 \ln m$ . Our fits strongly suggest that  $b_2 = 0$ , but we have been unable to find a rigorous analytic proof for this.

<sup>10</sup>Nevertheless this factor is, together with stability considerations, the reason why very small values of  $m$  cannot be reached.

L	Lüscher-Weisz		square	
	N=2	N=3	N=2	N=3
4	-0.021583156919197	-0.040661797276715	-0.020085367798214	-0.038849497567066
6	-0.019599952046337	-0.040131915718096	-0.018908108548525	-0.039452155788701
8	-0.018461969616743	-0.039274516206705	-0.018063236142647	-0.038897070150935
10	-0.017889223811184	-0.038810590019211	-0.017632105742522	-0.038571321147693
12	-0.017569982136115	-0.038547675035293	-0.017390811874792	-0.038382470120619
14	-0.017375041457146	-0.038385977285131	-0.017243147194999	-0.038265025623147
16	-0.017247595811504	-0.038279865552201	-0.017146489209537	-0.038187471819923
18	-0.017159817570292	-0.038206617036467	-0.017079864357098	-0.038133728057528
20	-0.017096835258186	-0.038153983851014	-0.017032035100709	-0.038095009442688
22	-0.017050132787137	-0.038114916592240	-0.016996555876247	-0.038066217556170
24	-0.017014554076415	-0.038085133354648	-0.016969519951262	-0.038044238228378
26	-0.016986831369867	-0.038061914071922	—	—

Table 5-2. Raw data for  $\bar{m}_{\Lambda^+}^{(1)}/m$ . For each entry the absolute accuracy is estimated to be  $(\frac{NL}{2\pi})^2 \cdot 10^{-14}$ .

We found the most accurate results by using a minimal set of points for the fit,  $L = 16 \dots 26$  for the Lüscher-Weisz action, and  $L = 14 \dots 24$  for the square action. The disadvantage of this approach is that the error estimate of the coefficients is necessarily ad hoc. As error estimate of  $a_i$  ( $i = 0, 1$ ) we typically used  $a_i - \bar{a}_i$ , where  $\bar{a}_i$  was obtained by dropping either  $a_3$  or  $b_3$  as a fit parameter. The justification for this procedure is that the coefficients  $a_3$  and  $b_3$  themselves can barely be determined from our data. As an a posteriori justification, the results below show the error estimate to be realistic whenever a comparison can be made to other results.

Using this method we obtain

$$a_0^{\text{LW}} = \begin{cases} -0.0168265791(7) & \text{for } N = 2 \\ -0.0379274963(15) & \text{for } N = 3, \end{cases} \quad (5.5.39)$$

$$a_1^{\text{LW}} = \begin{cases} -0.01100890(15) & \text{for } N = 2 \\ -0.0208015(5) & \text{for } N = 3, \end{cases} \quad (5.5.40)$$

$$a_0^{\text{sq}} = \begin{cases} -0.0168265790(15) & \text{for } N = 2 \\ -0.037927497(2) & \text{for } N = 3, \end{cases} \quad (5.5.41)$$

$$a_1^{\text{sq}} = \begin{cases} -0.0083763(3) & \text{for } N = 2 \\ -0.0154489(7) & \text{for } N = 3. \end{cases} \quad (5.5.42)$$

Note that the continuum coefficient  $a_0$  is independent of the action chosen, as it should be<sup>11</sup>. Moreover, using eq. (5.5.38) we see that the values for  $a_1^{\text{LW}}$  are consistent

<sup>11</sup>The discrepancy between our result for  $a_0(N = 3)$  and the value quoted by Lüscher and Weisz is due to a misprint [86] in ref. [27].

with the somewhat more accurate values<sup>12</sup> obtained by Lüscher and Weisz [27] (these are copied in eq. (5.4.52)). The values for  $a_1^{sq}$  agree with our static quark results, eq. (5.4.53).

We would like to report that for the LW action we have also done a computation of  $m_{\Lambda^-}^{(1)}$  (which couples to  $\pi_{33}(k_0; 0, m, 0)$ ), for  $N = 3$ . As expected, we found a continuum coefficient different from  $a_0(A^+)$ , namely  $a_0(A^-) = 0.0000679470(6)$ . A structural check on Symanzik improvement is that  $m_{\Lambda^-}^{(1)}$  can be improved simultaneously with  $m_{\Lambda^+}^{(1)}$  (and the static quark potential), as we found  $a_1^{LW}(A^-) = -0.0208011(3)$ . For  $N = 2$  we did a run for the Wilson action, and found  $a_0(A^-) = 0.00006029(14)$ . Our results for  $a_0(A^\pm)$  agree with those obtained from dimensional regularization by van Baal [87], which imply the analytic spin-splitting formulas

$$a_0(A^+) - a_0(A^-) = \begin{cases} -\frac{1}{6\pi^2} & \text{for } N = 2 \\ -\frac{3}{8\pi^2} & \text{for } N = 3. \end{cases} \quad (5.5.43)$$

### 5.5.3 Effective coupling constant

In ref. [26] Lüscher and Weisz define an effective coupling constant  $\lambda$  through

$$\sqrt{Z(\mathbf{k})Z(\mathbf{p})Z(\mathbf{q})} \sum_{j=1}^2 e_j \Gamma_3(k, 1, p, 2; q, j) = i\lambda f(k, p, q). \quad (5.5.44)$$

A number of new symbols show up. The external lines  $(k, 1)$  and  $(p, 2)$  correspond to on-shell A particles with positive spin, while  $(q, e)$  corresponds to an on-shell B particle with positive spin:

$$k = (iE(\mathbf{k}), \mathbf{k}); \quad \mathbf{k} = (0, m, ir), \quad (5.5.45)$$

$$p = (-iE(\mathbf{p}), \mathbf{p}); \quad \mathbf{p} = (m, 0, ir), \quad (5.5.46)$$

$$q = (0, \mathbf{q}); \quad \mathbf{q} = (-m, -m, -2ir), \quad (5.5.47)$$

$$e = (0; 1, -1, 0). \quad (5.5.48)$$

Note that these particles and polarizations are completely physical, as they can be created by Polyakov lines, see section 1.3. From now on  $k$ ,  $p$  and  $q$  always have the above special meaning. The energy  $E$  and wave function renormalization  $Z$  were defined in the previous subsection<sup>13</sup>. The value of  $r$  is defined by  $E(\mathbf{q}) = 0$  (we choose  $r > 0$ ). Note that at tree level, eq. (B.14) applies to all external propagators, with  $\varepsilon_\mu = \delta_{\mu,1}, \delta_{\mu,2}$  or  $e_\mu$ . In particular  $E(\mathbf{k}) = r = \frac{1}{2}\sqrt{2}m$  to leading order in  $m$  and  $g_0$ . Furthermore the equalities

$$E(\mathbf{p}) = E(\mathbf{k}), \quad Z(\mathbf{p}) = Z(\mathbf{k}) \quad (5.5.49)$$

<sup>12</sup>Lüscher and Weisz gained two digits by doing a mass calculation in three compact dimensions. Unfortunately this setting is unsuitable for the computation of the other improvement relation by use of scattering theory.

<sup>13</sup>Momenta implicitly label the particle type, and hence no additional label for the  $Z$  factors is necessary.

hold to all orders in perturbation theory, because  $\mathbf{k}$  and  $\mathbf{p}$  are related by symmetries of the discretized twisted tube [27].  $\Gamma_3(k, 1; p, 2; q, j)$  is the three-point function  $\langle \bar{A}_1(k) \bar{A}_2(p) \bar{A}_j(q) \rangle$ , but dropping the trivial factor  $\delta(k + p + q)$  and amputating the external lines in the usual way. Finally, the color factor  $f(k, p, q)$  is defined in appendix D. The coupling  $\lambda$  is a suitable parameter for on-shell improvement, because its square is proportional to the residue of the pole in the scattering amplitude for  $(A^+, A^+) \rightarrow (A^+, A^+)$ , appearing due to  $B^+$  exchange [26].

We expand

$$\lambda = g_0 \left\{ \lambda^{(0)} + g_0^2 \lambda^{(1)} + \mathcal{O}(g_0^4) \right\}. \quad (5.5.50)$$

The computation of the tree-level value  $\lambda^{(0)}$  simply amounts to substituting the momenta  $k$ ,  $p$  and  $q$  in the relevant expressions in appendices B, C and D (only  $V_{\mu\nu\rho}^{(i=2)}$  is not listed, but it can be looked up in ref. [78]). The result reads

$$Z_0(\mathbf{k}) = 1 - (\bar{c}_1 - c_2 + \frac{1}{12}) m^2 + \mathcal{O}(m^4), \quad (5.5.51)$$

$$Z_0(\mathbf{q}) = 1 + (\bar{c}_1 - c_2) m^2 + \mathcal{O}(m^4), \quad (5.5.52)$$

$$\begin{aligned} \sum_j e_j \Gamma_3^{\text{tree}}(k, 1; p, 2; q, j) &= -2f(k, p, q) g_0 \sum_j e_j V_{12j}(k, p, q) \\ &= if(k, p, q) \times (-8g_0 m) \left[ 1 - (4\bar{c}_1 - 3c_2 + \frac{7}{24}) m^2 + \mathcal{O}(m^4) \right]. \end{aligned} \quad (5.5.53)$$

It immediately follows that

$$\lambda^{(0)} = -8m \left\{ 1 - \frac{1}{2} m^2 \left[ 9(\bar{c}_1 - c_2 + \frac{1}{12}) + 2c_2 \right] + \mathcal{O}(m^4) \right\}, \quad (5.5.54)$$

and finally we see [27] why in eq. (5.2.4)  $c_2 = 0$  is the second on-shell improvement condition at tree-level. From now on we assume  $c_2 = 0$  and  $\bar{c}_1 - c_2 = -\frac{1}{12}$ .

At one-loop level the following formula can be derived from eq. (5.5.44) and the expressions given in the previous subsection (which can easily be extended to the  $B^+$  particle):

$$\begin{aligned} \frac{\lambda^{(1)}}{m} &= \left( 1 - \frac{1}{24} m^2 \right) \frac{\Gamma^{(1)}}{m} - \frac{4}{k_0} \frac{d}{dk_0} \pi_{11}(k) \Big|_{k_0=iE(\mathbf{k})} \\ &\quad - 2 \left( 1 - \frac{1}{12} m^2 \right) \frac{d^2}{dq_0^2} \left( \frac{1}{2} \sum_{i,j} e_i e_j \pi_{ij}(q) \right) \Big|_{q_0=0} + \mathcal{O}(m^4), \end{aligned} \quad (5.5.55)$$

where  $\Gamma^{(1)}$  is defined through

$$\sum_{j=1}^2 e_j \Gamma_3(k, 1; p, 2; q, j) = i g_0 f(k, p, q) \left\{ \Gamma^{(0)} + g_0^2 \Gamma^{(1)} + \mathcal{O}(g_0^4) \right\}. \quad (5.5.56)$$

In the above equations we may use tree-level expressions for the external momenta, i.e. we are free to redefine

$$\mathbf{k} = (iE_0(\mathbf{k}), \mathbf{k}); \quad \mathbf{k} = (0, m, i r_0), \quad (5.5.57)$$

$$\mathbf{p} = (-iE_0(\mathbf{p}), \mathbf{p}); \quad \mathbf{p} = (m, 0, i r_0), \quad (5.5.58)$$

$$\mathbf{q} = (0, \mathbf{q}); \quad \mathbf{q} = (-m, -m, -2i r_0), \quad (5.5.59)$$

where  $r_0 > 0$  is the solution to  $E_0(\mathbf{q}) = 0$ . This redefinition brings about only  $g_0^4$  corrections to eq. (5.5.50). One should however be aware of the fact that in the derivation of eq. (5.5.54) we used eqs. (5.5.57)–(5.5.59) rather than eqs. (5.5.45)–(5.5.47). A priori this can cause corrections to  $\lambda^{(1)}$ . However, we have checked that all such corrections are at least of order  $m^4$ . For example in the second term of eq. (5.5.24),  $E_0$  and (at least for a physical polarization)  $E_1$  are of order  $m$ , while  $((2k_0)^{-1}d/dk_0)^2 d_{[0]}^{-1}$  is of order  $m^2$  due to tree-level improvement.

In terms of Feynman diagrams,  $\Gamma^{(1)}$  is represented by

$$(5.5.60)$$

Correspondingly we write

$$\Gamma^{(1)} = \Gamma^V + \Gamma^{\text{gh}1} + \Gamma^{\text{gh}2} + \Gamma^{\text{gh}3} + \Gamma^{\text{gh}4} + \Gamma^{\text{gh}5} + \Gamma^{V_1} + \Gamma^{V_2} + \Gamma^{V_3} + \Gamma^{W_1} + \Gamma^{W_2} + \Gamma^{W_3} + \Gamma^{V_5} + \Gamma^{V_6}, \quad (5.5.61)$$

where

$$\Gamma^V = 2i \sum_j e_j \sum_i c_i^V V_{12j}^{(1)}(k, p, q), \quad (5.5.62)$$

$$\Gamma^{\text{gh}1} = 0, \quad (5.5.63)$$

$$\Gamma^{\text{gh}2} = \frac{1}{12} \oint_{k', k''} \delta(k' + k'' - k) G_{k, p, q}^{k', k'', k'} \frac{\tilde{k}_2' \tilde{k}_2'' (k' - k'')_1}{i \tilde{k}_2'^2 \tilde{k}_2''^2}, \quad (5.5.64)$$

$$\Gamma^{gh3} = \Gamma^{gh2}, \quad (5.5.65)$$

$$\Gamma^{gh4} = - \sum_{k', k'', k'''} \delta(k + k' - k''') \delta(p + k'' - k') G_{k,p,q}^{k', k'', k'''} \frac{\hat{k}'_1 \hat{k}''_2 c_1''' c_2' (\hat{k}'''_1 c_1' - \hat{k}''_2 c_2''')}{k'^2 k''^2 k'''^2}, \quad (5.5.66)$$

$$\Gamma^{gh5} = \Gamma^{gh4}, \quad (5.5.67)$$

$$\Gamma^{V_4^1} = \frac{1}{6} i \sum_j e_j \sum_{k', k''} \delta(k' + k'' - q) D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') V_{\lambda\rho j}(-k', -k'', q) \\ \times \{g_-^2(k', k'') \bar{V}_{12\lambda'\rho'}(k, p, k', k'') - 2G_{k,p,q}^{k', k'', k'} \bar{V}_{1\lambda'2\rho'}(k, k', p, k'')\}, \quad (5.5.68)$$

$$\Gamma^{V_4^2} = -\frac{1}{6} i \sum_j e_j \sum_{k', k''} \delta(k' + k'' - k) D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') V_{\lambda\rho 1}(-k', -k'', k) \\ \times \{g_-^2(k', k'') \bar{V}_{j2\lambda'\rho'}(q, p, k', k'') + 2G_{k,p,q}^{k', k'', k'} \bar{V}_{j\lambda'2\rho'}(q, k', p, k'')\}, \quad (5.5.69)$$

$$\Gamma^{V_4^3} = \Gamma^{V_4^2}, \quad (5.5.70)$$

$$\Gamma^{W^1} = \frac{1}{12} i \sum_j e_j \sum_{k', k''} \delta(k' + k'' - q) D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') V_{\lambda\rho j}(-k', -k'', q) \frac{g_-(k', k'')}{g_-(k, p)} \\ \times \{g_+(k', k'') g_+(k, p) + (k' \leftrightarrow k) + (k' \leftrightarrow p)\} W_{12\lambda'\rho'}(k, p, k', k''), \quad (5.5.71)$$

$$\Gamma^{W^2} = \frac{1}{12} i \sum_j e_j \sum_{k', k''} \delta(k' + k'' - k) D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') V_{\lambda\rho 1}(-k', -k'', k) \frac{g_-(k', k'')}{g_-(k, p)} \\ \times \{g_+(k', k'') g_+(q, p) + (k' \leftrightarrow q) + (k' \leftrightarrow p)\} W_{j2\lambda'\rho'}(q, p, k', k''), \quad (5.5.72)$$

$$\Gamma^{W^3} = \Gamma^{W^2}, \quad (5.5.73)$$

$$\Gamma^{V_5} = -i \sum_j e_j \sum_{k', k'', k'''} \delta(k + k' - k''') \delta(p + k'' - k') D_{\lambda\lambda'}(k') D_{\rho\rho'}(k'') D_{\tau\tau'}(k''') \\ \times G_{k,p,q}^{k', k'', k'''} V_{\lambda'\tau 1}(k', -k''', k) V_{\rho'\lambda 2}(k'', -k', p) V_{\tau'\rho j}(k''', -k'', q), \quad (5.5.74)$$

$$\Gamma^{V_6} = \frac{1}{24} i \sum_j e_j \sum_{k'} D_{\lambda\rho}(k') \chi_{k'} \{ [V_{12j\lambda\rho}(k, p, q, k', -k') + 2 \text{ cyclic perms}] \\ + \frac{1}{g_-(k,p)} [g_-(k, p + 2k') V_{1\lambda 2j\rho}(k, k', p, q, -k') + 2 \text{ cyclic p.}] \}. \quad (5.5.75)$$

In the latter equation, the cyclic permutations act on  $(k, 1)$ ,  $(p, 2)$  and  $(q, j)$ . Furthermore,  $g_-$  is defined in eq. (5.5.36) and

$$g_+(k_1, k_2) \equiv 2 \cos \left( \frac{\pi}{N} (k_1, k_2) \right), \quad (5.5.76)$$

$$G_{k_1, k_2, k_3}^{k_4, k_5, k_6} \equiv \frac{g_-(k_1, k_4) g_-(k_2, k_5) g_-(k_3, k_6)}{g_-(k_1, k_2)}, \quad (5.5.77)$$

$$\bar{V}_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4) \equiv V_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4) - V_{\mu_2 \mu_1 \mu_3 \mu_4}(k_2, k_1, k_3, k_4). \quad (5.5.78)$$

The equality of various Feynman diagrams is due to lattice symmetries and properties of  $k, p$  and  $q$ . We checked that in leading order in  $m$  the expressions reduce to their infinite volume counterparts, cf. the discussion below eq. (5.5.36). Numerically we found that the  $\Gamma^W$  contributions are at least of order  $m^4$ .

L	Lüscher-Weisz		square	
	N=2	N=3	N=2	N=3
4	-0.78341711803619	-1.47968692446631	-0.79859872384707	-1.51787874584542
6	-0.99849061664701	-1.78876943596789	-1.01681082402891	-1.81613837864479
8	-1.13963010017710	-1.98424388671692	-1.15382728105870	-2.00463419279200
10	-1.24091978060661	-2.12599811096867	-1.25207004886790	-2.14258204874697
12	-1.31941153232092	-2.23737552735037	-1.32855710935494	-2.25172969936027
14	-1.38345219535201	-2.32928849922146	-1.39125843561358	-2.34224049339430
16	-1.43757858326173	-2.40765025540042	-1.44445943234635	-2.41966813031578
18	-1.48449250619699	-2.47601715882150	-1.49071210672691	-2.48738335434952
20	-1.52592324035698	-2.53669577194007	-1.53165594879818	-2.54759007845263

Table 5-3. Raw data for  $\bar{\lambda}^{(1)}/m$ . For each entry the absolute accuracy is estimated to be  $10^{-13}$ . Note that the inexact formula (5.5.55) was used, so that the data has systematic deviations of order  $m^4$ .

The analytic expressions for the second and third terms in eq. (5.5.55) can be found by differentiating eqs. (5.5.29)–(5.5.35) (the differentiation can be brought over to the integrands without problem). We checked the corresponding computer programs against our old programs for  $\pi_{\mu\nu}(\vec{k})$  by running the latter for near values of  $k_0$ .

The diagrams proportional to  $c_i'$  can be calculated analytically. We separate their total contribution (which can be read off from eq. (5.5.54)):

$$\frac{\lambda^{(1)}}{m} = \frac{\bar{\lambda}^{(1)}}{m} + 4m^2 [9(c_1' - c_2') + 2c_2'] + \mathcal{O}(m^4). \quad (5.5.79)$$

Like  $\bar{m}_{\Lambda^+}^{(1)}$ ,  $\bar{\lambda}^{(1)}$  is to be evaluated numerically for a number of lattice sizes. The remarks in the previous subsection concerning the integration routine are also valid in the present case. Thus the translation of the analytic expressions to computer programs is straightforward. The resulting data is listed in table 5-3.

The data is expected to be of the form [27] (with new coefficients  $a_i$  and  $b_i$ )

$$\frac{\bar{\lambda}^{(1)}}{m} = a_0 + b_0 \ln m + a_1 m^2 + a_2 m^4 + b_2 m^4 \ln m + a_3 m^6 + b_3 m^6 \ln m + \mathcal{O}(m^8). \quad (5.5.80)$$

Indeed it can be checked that our data is consistent with the absence of odd powers in  $m$ . Also we checked the absence of an  $m^2 \ln m$  term, expected due to tree-level improvement. Furthermore, since  $\lim_{m \rightarrow 0} (\lambda/m)_{\text{tree level}} = -8g_0$  and  $\lambda$  is a renormalizable parameter,  $b_0$  should equal

$$b_0 = 8\beta_0 = \frac{11N}{6\pi^2}. \quad (5.5.81)$$

The fit (5.5.80) (dropping  $a_3$  or  $b_3$ ) to our data for  $L = 10 \dots 20$  reproduces this value to six digits. This a non-trivial check against programming errors (and other errors)<sup>14</sup>. Moreover, by fixing  $b_0$  to have the exact value (5.5.81) the other coefficients can be fitted to a higher accuracy. In this way we obtain

$$a_0^{\text{LW}} = \begin{cases} -0.84832346(3) & \text{for } N = 2 \\ -1.28773532(5) & \text{for } N = 3, \end{cases} \quad (5.5.82)$$

$$a_1^{\text{LW}} = \begin{cases} 0.419861(6) & \text{for } N = 2 \\ 0.78417(3) & \text{for } N = 3, \end{cases} \quad (5.5.83)$$

$$a_0^{\text{sq}} = \begin{cases} -0.85183887(3) & \text{for } N = 2 \\ -1.29656105(4) & \text{for } N = 3, \end{cases} \quad (5.5.84)$$

$$a_1^{\text{sq}} = \begin{cases} 0.324745(5) & \text{for } N = 2 \\ 0.59095(2) & \text{for } N = 3, \end{cases} \quad (5.5.85)$$

where the error-estimate procedure described in the previous subsection was used.

The LW coefficients agree with the ones obtained by Lüscher and Weisz [27]:  $a_0^{\text{LW}}(N = 2) = -0.8483231(3)$ ,  $a_0^{\text{LW}}(N = 3) = -1.2877352(1)$ ,  $a_1^{\text{LW}}(N = 2) = 0.41988(3)$ ,  $a_1^{\text{LW}}(N = 3) = 0.78412(5)$ . The  $a_0$  coefficient for the square and LW actions differ. This is a renormalization effect. Like in subsection 5.4.1 one can see that the  $a_0$  coefficient is related to the Lambda parameter:

$$\frac{\Lambda^*}{\Lambda} = e^{\frac{1}{t_0}(a_0(t;)) - a_0(t;)} \quad (5.5.86)$$

We thus find

$$\frac{\Lambda_{\text{LW}}}{\Lambda_{\text{sq}}} = \begin{cases} 1.0095074(2) & \text{for } N = 2 \\ 1.0159636(2) & \text{for } N = 3, \end{cases} \quad (5.5.87)$$

in complete agreement with eq. (5.4.51). Incidentally, for  $N = 2$  we also did a run for the Wilson action and found

$$a_0^{\text{W}}(N = 2) = -1.37530949(10), \quad (5.5.88)$$

so that

$$\frac{\Lambda_{\text{LW}}}{\Lambda_{\text{W}}} = 4.1308934(14) \quad \text{for } N = 2, \quad (5.5.89)$$

$$\frac{\Lambda_{\text{sq}}}{\Lambda_{\text{W}}} = 4.0919894(14) \quad \text{for } N = 2. \quad (5.5.90)$$

<sup>14</sup>Useful intermediate checks can be obtained by extracting the coefficients of the  $m^0 \ln m$  terms for individual diagrams, and comparing them to the  $1/(4-d)$  poles of their dimensionally regularized continuum counterparts.

After renormalization one reads off the improvement condition:

$$4[9(\bar{c}'_1 - \bar{c}'_2) + 2\bar{c}'_3] = -a_1. \quad (5.5.91)$$

Hence together with the results from the previous subsection we have completely determined the one-loop Symanzik coefficients for the square (and Lüscher-Weisz) action.

We are convinced of the correctness of our new result (5.5.85) because of the many internal checks in our computation. In particular, it is extremely implausible that an expansion giving the correct values for the leading coefficients  $b_0$  and  $a_0$  (as verified from the Lambda ratios) as well as for the subleading coefficient  $b_1$ , would produce an incorrect value for the other subleading coefficient  $a_1$ . The fact that for the LW action all results by Lüscher and Weisz have been reproduced, confirms our confidence to a large extent.

## 5.6 Summary

In this chapter we computed Lambda ratios, one-loop Symanzik coefficients and the tadpole parameter for the Lüscher-Weisz and square actions. The results are summarized in table 5-4. We remind the reader that coefficient combinations other than  $\bar{c}_1(g_0^2) \equiv c_1(g_0^2) + 4c_4(g_0^2)$  and  $c_2(g_0^2)$  are unimportant because only  $\bar{c}_1(g_0^2)$  and  $c_2(g_0^2)$  couple to the  $a^2$  corrections of any on-shell quantities (see section 5.2).

For the Lüscher-Weisz action we found complete agreement with previous calculations [27, 78]. In particular we reproduced the values for the Lambda parameter and the one-loop coefficients, as well as for  $l(1, 1)$  and  $l(1, 2)$  (which are related to the tadpole parameter  $u_0$ ). Especially the agreement with ref. [27] is non-trivial because we used covariant instead of Coulomb gauge fixing. For the square action we computed the combination  $\bar{c}'_1 - \bar{c}'_2$  and the Lambda parameter in two independent ways, namely using the static quark potential and finite volume spectroscopy. The Lambda parameter was also computed in ref. [84] from a background field method. The agreement between all results leaves us with no doubt that our values for  $(\bar{c}'_1 - \bar{c}'_2)_{\text{square}}$  and  $\Lambda_{\text{square}}$  are correct. We are also convinced of the correctness of our result for the combination  $(9(\bar{c}'_1 - \bar{c}'_2) + 2\bar{c}'_3)_{\text{square}}$ , because it was found from the same numerical data that gave correct results for three verifiable coefficients.

We conclude with testing how well the tadpole correction [29] to the SU(3) tree-level square action predicts the one-loop correction. For  $\bar{c}'_2$  the prediction is  $c_2/u_0^2 = 0$ , and indeed in table 5-4  $\bar{c}'_2$  is much smaller than  $\bar{c}'_1$ . In the  $(c_1(g_0^2), c_4(g_0^2))$  sector one has to keep in mind that, to  $\mathcal{O}(a^2)$ ,  $c_4(g_0^2)$  can be chosen freely. The relevant test is comparing

$$\frac{\bar{c}'_1(g_0^2)}{\bar{c}'_0(g_0^2)} = -\frac{1}{20}(1 + 0.1217g_0^2 + \mathcal{O}(g_0^4)) \quad (5.6.1)$$

to

$$\frac{c_1 u_0^{-2} + 4c_4 u_0^{-4}}{c_0 - 16c_4 u_0^{-4}} = -\frac{1}{20}(1 + 0.0957g_0^2 + \mathcal{O}(g_0^4)). \quad (5.6.2)$$

	Lüscher-Weisz	square
$\tilde{c}_0$	0.135160(13) ( $N = 2$ )	0.113417(11) ( $N = 2$ )
	0.23709(6) ( $N = 3$ )	0.19320(4) ( $N = 3$ )
$\tilde{c}_1$	-0.0139519(8) ( $N = 2$ )	-0.0112766(7) ( $N = 2$ )
	-0.025218(4) ( $N = 3$ )	-0.019799(2) ( $N = 3$ )
$\tilde{c}_2$	-0.0029431(8) ( $N = 2$ )	-0.0029005(7) ( $N = 2$ )
	-0.004418(4) ( $N = 3$ )	-0.004351(2) ( $N = 3$ )
$\Lambda/\Lambda_{\text{Wilson}}$	4.1308935(3) ( $N = 2$ )	4.0919901(2) ( $N = 2$ )
	5.2921038(3) ( $N = 3$ )	5.2089503(2) ( $N = 3$ )
$\mathbb{I}(1, 1)$	0.366262680(2)	0.3587838551(1)
$\mathbb{I}(1, 2)$	0.662626785(2)	0.6542934512(1)
$\mathbb{I}(2, 2)$	1.098143594(2)	1.0887235337(1)

Table 5-4. One-loop improvement coefficients, Lambda parameter ratios and one-loop expectation values of small Wilson loops for the Lüscher-Weisz and square Symanzik actions. For quantities that were extracted both from the static quark potential and from the twisted spectroscopy, we used the most accurate result. In particular we made use of ref. [27] for the combination  $\tilde{c}_1 - \tilde{c}_2$ , cf. subsection 5.4.2, which we verified to somewhat lower accuracy in subsection 5.5.2. For the definition of  $\mathbb{I}(L, T)$  and its relation to the tadpole parameter, see eqs. (5.4.2), (5.4.22) and (5.4.24).

Here  $u_0 = 1 - 0.3588g_0^2/6$  was taken from table 5-4. It follows that the tadpole prediction captures 79% of the one-loop correction, a result similar to the 76% found for the Lüscher-Weisz action. (For SU(2) one finds 80% for both actions). In this sense tadpole improvement performs well, *consistently* for both Symanzik improved actions.

For the square action one might wish to consider the ratios  $c_1(g_0^2)/c_0(g_0^2)$  and  $c_4(g_0^2)/c_0(g_0^2)$  separately. While for  $c_4(N = 3) = 0.003058$ , satisfying  $c_4(g_0^2)c_0(g_0^2) = (c_1(g_0^2))^2$  to one-loop order, the tadpole prediction is off by 21% in both ratios, for  $c_4 = 0.002401$  the deviations are only 11%. For SU(2) the deviations are 20% for  $c_4 = 0.001718$  and 10% for  $c_4 = 0.001377$ .

## Appendix A: Structure of the Feynman rules

In this appendix we expand the total action, eq. (5.3.5), in powers of  $g_0$ , largely adopting the notation of refs. [27, 78, 79]. We use Fourier space:

$$A_\mu(x) = g_0 \sum_{b=1}^{N^2-1} \int_k e^{ik(x+\frac{1}{2}a\mu)} \bar{A}_\mu^b(k) T^b \quad (\text{A.1})$$

where

$$\int_k \equiv \prod_{\mu=0}^3 \left( \int_{-\pi/a}^{\pi/a} \frac{dk_\mu}{2\pi} \right). \quad (\text{A.2})$$

It should be noted that the above definitions hold for the infinite-volume case. The modifications for the finite twisted volume are discussed in appendix D. From now on, summations over  $SU(N)$  and Lorentz indices will be implicit.

The expansion of the measure and ghost actions can be found in refs. [27, 78]. For completeness we copy the results<sup>15</sup>:

$$S_{\text{measure}} = \frac{N}{24} g_0^2 \delta_{\mu_1 \mu_2} \delta_{a_1 a_2} \frac{1}{a^2} \int_{k_1, k_2} (2\pi)^4 \delta(k_1 + k_2) \bar{A}_{\mu_1}^{a_1}(k_1) \bar{A}_{\mu_2}^{a_2}(k_2) + \mathcal{O}(g_0^4), \quad (\text{A.3})$$

$$\begin{aligned} S_{\text{ghost}} = & \int_{k_1, k_2} \bar{c}^{a_1}(k_1) \bar{c}^{a_2}(k_2) \left[ (2\pi)^4 \delta(k_1 + k_2) \delta_{a_1 a_2} \hat{k}_1^2 + i g_0 \int_{a_1 a_2 a_3} \right. \\ & \times \int_{k_3} (2\pi)^4 \delta(k_1 + k_2 + k_3) \bar{A}_{\mu}^{a_3}(k_3) \hat{k}_{1\mu} c_{2\mu} + \frac{1}{12} g_0^2 \delta_{\mu_3 \mu_4} f_{a_1 a_3 c} f_{a_2 a_4 c} \\ & \left. \times a^2 \int_{k_3, k_4} (2\pi)^4 \delta(k_1 + k_2 + k_3 + k_4) \bar{A}_{\mu_3}^{a_3}(k_3) \bar{A}_{\mu_4}^{a_4}(k_4) \hat{k}_{1\mu_3} \hat{k}_{2\mu_4} + \mathcal{O}(g_0^4) \right]. \end{aligned} \quad (\text{A.4})$$

Here  $\bar{c}^a$ ,  $\bar{c}^{\bar{a}}$  are the Fourier-transformed ghost fields. Also we adopted the conventional notation  $\hat{k}^2 = \sum_{\mu=0}^3 \hat{k}_\mu^2$  and

$$\begin{aligned} \hat{k}_\mu &= \frac{2}{a} \sin\left(\frac{a}{2} k_\mu\right), \\ c_\mu &= \cos\left(\frac{a}{2} k_\mu\right). \end{aligned} \quad (\text{A.5})$$

The difference between  $c_\mu$  and the action coefficients  $c_i$  should always be clear from the context.

The remaining part of  $S_{\text{total}}/g_0^2$  is expanded as follows:

$$\frac{1}{g_0} (S(\{c_i(g_0^2)\}) + S_{\text{gr}}) = \sum_{n=2}^{\infty} \frac{g_0^{n-2}}{n!} S_n(\{c_i(g_0^2)\}). \quad (\text{A.6})$$

(Unlike in the continuum this expansion does not truncate at  $n=4$ ).  $S_n(\{c_i(g_0^2)\})$  is linear in  $c_i(g_0^2)$  and therefore we have

<sup>15</sup>There is a slight overall difference with ref. [78] because we explicitly keep a factor  $(-1/g_0^2)$  in the path integral, eq. (5.3.6), instead of absorbing it into the action.

$$S_n(\{c_i(g_0^2)\}) = S_n(\{c_i\}) + g_0^2(S_n(\{c'_i\}) - S_n(\{0\})) + \mathcal{O}(g_0^4). \quad (\text{A.7})$$

In sections 5.4 and 5.5 the terms  $g_0^2(S_n(\{c'_i\}) - S_n(\{0\}))$  are treated as insertions.

Only  $S_2$ ,  $S_3$ ,  $S_4$  and  $S_5$  are needed in our calculations. These can be written in the following way:

$$S_2(\{c_i\}) = \delta_{a_1 a_2} \int_{k_1, k_2} (2\pi)^4 \delta(k_1 + k_2) \bar{A}_{\mu_1}^{a_1}(k_1) \bar{A}_{\mu_2}^{a_2}(k_2) (D^{-1})_{\mu_1 \mu_2}(k_1), \quad (\text{A.8})$$

$$S_3(\{c_i\}) = f_{a_1 a_2 a_3} \int_{k_1, k_2, k_3} (2\pi)^4 \delta(k_1 + k_2 + k_3) \times \bar{A}_{\mu_1}^{a_1}(k_1) \bar{A}_{\mu_2}^{a_2}(k_2) \bar{A}_{\mu_3}^{a_3}(k_3) V_{\mu_1 \mu_2 \mu_3}(k_1, k_2, k_3), \quad (\text{A.9})$$

$$S_4(\{c_i\}) = \int_{k_1, k_2, k_3, k_4} (2\pi)^4 \delta(k_1 + k_2 + k_3 + k_4) \bar{A}_{\mu_1}^{a_1}(k_1) \bar{A}_{\mu_2}^{a_2}(k_2) \bar{A}_{\mu_3}^{a_3}(k_3) \bar{A}_{\mu_4}^{a_4}(k_4) \times [f_{a_1 a_2 e} f_{a_3 a_4 e} (V_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4) - V_{\mu_2 \mu_1 \mu_3 \mu_4}(k_2, k_1, k_3, k_4)) - S_{a_1 a_2 a_3 a_4} W_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4)], \quad (\text{A.10})$$

$$S_5(\{c_i\}) = C_{a_1 a_2 a_3 a_4 a_5} \int_{k_1, k_2, k_3, k_4, k_5} (2\pi)^4 \delta(k_1 + k_2 + k_3 + k_4 + k_5) \times \bar{A}_{\mu_1}^{a_1}(k_1) \cdots \bar{A}_{\mu_5}^{a_5}(k_5) V_{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5}(k_1, k_2, k_3, k_4, k_5), \quad (\text{A.11})$$

where we defined

$$\begin{aligned} V_{\mu_1 \mu_2 \mu_3} &= \sum_i c_i V_{\mu_1 \mu_2 \mu_3}^{(i)}, \\ V_{\mu_1 \mu_2 \mu_3 \mu_4} &= \sum_i c_i V_{\mu_1 \mu_2 \mu_3 \mu_4}^{(i)}, \\ W_{\mu_1 \mu_2 \mu_3 \mu_4} &= \sum_i c_i W_{\mu_1 \mu_2 \mu_3 \mu_4}^{(i)}, \\ V_{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5} &= \sum_i c_i V_{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5}^{(i)}. \end{aligned} \quad (\text{A.12})$$

The definitions of the color factors appearing in eqs. (A.10), (A.11) are:

$$S_{abcd} = \frac{1}{24} \text{Tr} (T^a T^b T^c T^d + 23 \text{ permutations}), \quad (\text{A.13})$$

$$C_{abcde} = \text{Tr} (T^a T^b T^c T^d T^e - T^e T^d T^c T^b T^a). \quad (\text{A.14})$$

For our purposes it is not necessary to work out these factors for general values of the indices. Note that  $W_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4)$  is completely symmetric under permutations of  $1 \cdots 4$  and vanishes in the continuum limit.

The Feynman propagator  $D_{\mu\nu}$  and its inverse are discussed in appendix B, while in appendix C the vertices are given. The following parametrization, valid as long as only planar Wilson loops are included in the action, is convenient:

$$V_{\mu_1\mu_2\mu_3}(k_1, k_2, k_3) = f_{\mu_1}^{(3)}(k_1, k_2, k_3)\delta_{\mu_1\mu_2\mu_3} + [g_{\mu_1\mu_3}^{(3)}(k_1, k_2, k_3)\delta_{\mu_1\mu_2} + 2 \text{ cyclic permutations}], \quad (\text{A.15})$$

$$V_{\mu_1\mu_2\mu_3\mu_4}(k_1, k_2, k_3, k_4) = f_{\mu_1}^{(4)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2\mu_3\mu_4} + [g_{\mu_1\mu_4}^{(4)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2\mu_3} + 3 \text{ cyclic perms}] + h_{\mu_1\mu_2}^{(4)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_3}\delta_{\mu_2\mu_4} + [h_{\mu_1\mu_3}^{(4)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2}\delta_{\mu_3\mu_4} + 1 \text{ cyclic perm}], \quad (\text{A.16})$$

$$W_{\mu_1\mu_2\mu_3\mu_4}(k_1, k_2, k_3, k_4) = f_{\mu_1}^{(W)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2\mu_3\mu_4} + [g_{\mu_1\mu_4}^{(W)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2\mu_3} + 3 \text{ cyclic perms}] + [h_{\mu_1\mu_3}^{(W)}(k_1, k_2, k_3, k_4)\delta_{\mu_1\mu_2}\delta_{\mu_3\mu_4} + (2 \leftrightarrow 3) + (2 \leftrightarrow 4)], \quad (\text{A.17})$$

$$V_{\mu_1\mu_2\mu_3\mu_4\mu_5}(k_1, k_2, k_3, k_4, k_5) = f_{\mu_1}^{(5)}(k_1, k_2, k_3, k_4, k_5)\delta_{\mu_1\mu_2\mu_3\mu_4\mu_5} + [g_{\mu_1\mu_5}^{(5)}(k_1, k_2, k_3, k_4, k_5)\delta_{\mu_1\mu_2\mu_3\mu_4} + 4 \text{ cyclic perms}] + [h_{\mu_1\mu_2}^{(5)}(k_1, k_2, k_3, k_4, k_5)\delta_{\mu_1\mu_3\mu_5}\delta_{\mu_2\mu_4} + 4 \text{ cyclic perms}] + [h_{\mu_1\mu_4}^{(5)}(k_1, k_2, k_3, k_4, k_5)\delta_{\mu_1\mu_2\mu_3}\delta_{\mu_4\mu_5} + 4 \text{ cyclic perms}], \quad (\text{A.18})$$

where  $\delta_{\mu_1 \dots \mu_n} \equiv \delta_{\mu_1\mu_2} \dots \delta_{\mu_{n-1}\mu_n}$ . Permutations act simultaneously on Lorentz indices and momenta. From considerations given in ref. [27] it follows that vertices  $V_{\mu_1 \dots \mu_n}$ ,  $n$  even, are real and invariant under inversion of all momenta. For  $n$  odd they are imaginary and odd under inversion. Moreover, the components have the following properties with respect to permutations:

$$\begin{aligned} f_{\mu}^{(3)}(k_1, k_2, k_3) &= -f_{\mu}^{(3)}(k_3, k_2, k_1) = f_{\mu}^{(3)}(k_2, k_3, k_1), \\ g_{\mu\nu}^{(3)}(k_1, k_2, k_3) &= -g_{\mu\nu}^{(3)}(k_2, k_1, k_3), \\ f_{\mu}^{(4)}(k_1, k_2, k_3, k_4) &= f_{\mu}^{(4)}(k_4, k_3, k_2, k_1) = f_{\mu}^{(4)}(k_2, k_3, k_4, k_1), \\ g_{\mu\nu}^{(4)}(k_1, k_2, k_3, k_4) &= g_{\mu\nu}^{(4)}(k_3, k_2, k_1, k_4), \\ h_{\mu\nu}^{(4)}(k_1, k_2, k_3, k_4) &= h_{\mu\nu}^{(4)}(k_4, k_3, k_2, k_1) = h_{\nu\mu}^{(4)}(k_2, k_3, k_4, k_1), \\ h_{\mu\nu}^{(4)}(k_1, k_2, k_3, k_4) &= h_{\nu\mu}^{(4)}(k_4, k_3, k_2, k_1) = h_{\mu\nu}^{(4)}(k_3, k_4, k_1, k_2), \\ f_{\mu}^{(W)}(k_1, k_2, k_3, k_4) &= f_{\mu}^{(W)}(k_{P(1)}, k_{P(2)}, k_{P(3)}, k_{P(4)}) \text{ for all perms } P, \\ g_{\mu\nu}^{(W)}(k_1, k_2, k_3, k_4) &= g_{\mu\nu}^{(W)}(k_3, k_2, k_1, k_4) = g_{\mu\nu}^{(W)}(k_2, k_3, k_1, k_4), \\ h_{\mu\nu}^{(W)}(k_1, k_2, k_3, k_4) &= h_{\mu\nu}^{(W)}(k_2, k_1, k_3, k_4) = h_{\mu\nu}^{(W)}(k_1, k_2, k_4, k_3) = \\ &h_{\nu\mu}^{(W)}(k_3, k_4, k_1, k_2), \\ f_{\mu}^{(5)}(k_1, k_2, k_3, k_4, k_5) &= -f_{\mu}^{(5)}(k_5, k_4, k_3, k_2, k_1) = f_{\mu}^{(5)}(k_2, k_3, k_4, k_5, k_1), \\ g_{\mu\nu}^{(5)}(k_1, k_2, k_3, k_4, k_5) &= -g_{\mu\nu}^{(5)}(k_4, k_3, k_2, k_1, k_5), \\ h_{\mu\nu}^{(5)}(k_1, k_2, k_3, k_4, k_5) &= -h_{\mu\nu}^{(5)}(k_5, k_4, k_3, k_2, k_1), \\ h_{\mu\nu}^{(5)}(k_1, k_2, k_3, k_4, k_5) &= -h_{\mu\nu}^{(5)}(k_3, k_2, k_1, k_5, k_4). \end{aligned} \quad (\text{A.19})$$

## Appendix B: Propagator

Following the notation of ref. [79], the inverse propagator as defined in eq. (A.8) reads

$$(D^{-1})_{\mu\nu}(k) = \left[ \sum_{\rho} q_{\mu\rho}(k) \hat{k}_{\rho}^2 \right] \delta_{\mu\nu} - \left[ q_{\mu\nu}(k) - \frac{1}{\alpha} \right] \hat{k}_{\mu} \hat{k}_{\nu}. \quad (\text{B.1})$$

For the action in eq. (4.2.1) the tensor  $q_{\mu\nu}$  equals

$$\begin{aligned} q_{\mu\nu}(k) &= (1 - \delta_{\mu\nu}) s_{\mu\nu}(k), \\ s_{\mu\nu}(k) &= (c_0 + 8c_1 + 8c_2 + 16c_4) - a^2(c_1 - c_2 + 4c_4)(\hat{k}_{\mu}^2 + \hat{k}_{\nu}^2) \\ &\quad - a^2 c_2 \hat{k}^2 + a^4 c_4 \hat{k}_{\mu}^2 \hat{k}_{\nu}^2. \end{aligned} \quad (\text{B.2})$$

(In view of the discussion in section 5.2 we do not include  $c_3$ ). Note that, by definition, this tensor factorizes for the square action ( $c_2 = 0$ ,  $c_0 c_4 = c_1^2$ ).

The inverse of eq. (B.1) reads

$$D_{\mu\nu}(k) = \frac{1}{(\hat{k}^2)^2} \left\{ \left[ \sum_{\rho} A_{\mu\rho}(k) \hat{k}_{\rho}^2 \right] \delta_{\mu\nu} - [A_{\mu\nu}(k) - \alpha] \hat{k}_{\mu} \hat{k}_{\nu} \right\}, \quad (\text{B.3})$$

with  $A_{\mu\nu}$  satisfying

$$A_{\mu\mu}(k) = 0 \quad (\forall \mu), \quad A_{\mu\nu}(k) = A_{\nu\mu}(k) = A_{\mu\nu}(-k), \quad \lim_{a \rightarrow 0} A_{\mu\nu}(k) \stackrel{\mu \neq \nu}{=} \frac{1}{s_{\mu\nu}(0)}. \quad (\text{B.4})$$

In ref. [79] the general form of  $A_{\mu\nu}$  in terms of  $q_{\mu\nu}$  can be found. Using the notation

$$\hat{k}^{(n)} = \sum_{\rho=0}^3 \hat{k}_{\rho}^n, \quad (\text{B.5})$$

$$\tilde{\mu}, \tilde{\nu} : \text{such that } \mu \neq \nu \neq \tilde{\mu} \neq \tilde{\nu}, \quad (\text{B.6})$$

$$P_{\mu\nu} = \hat{k}_{\tilde{\mu}}^2 \hat{k}_{\tilde{\nu}}^2, \quad (\text{B.7})$$

$$Q_{\mu\nu} = \hat{k}_{\tilde{\mu}}^2 + \hat{k}_{\tilde{\nu}}^2, \quad (\text{B.8})$$

we here give the explicit formulas for the Wilson (W), Lüscher-Weisz (LW) [78, 79] and square (sq) cases:

$$A_{\mu\nu}^{(W)}(k) = 1 - \delta_{\mu\nu}, \quad (\text{B.9})$$

$$A_{\mu\nu}^{(LW)}(k) = \frac{1 - \delta_{\mu\nu}}{\Delta^{(LW)}} \left( \hat{k}^2 + \frac{1}{12} (\hat{k}^{(4)} + \hat{k}^2 \hat{k}_{\tilde{\mu}}^2) \right) \left( \hat{k}^2 + \frac{1}{12} (\hat{k}^{(4)} + \hat{k}^2 \hat{k}_{\tilde{\nu}}^2) \right), \quad (\text{B.10})$$

$$\begin{aligned} \Delta^{(LW)} &= \left( \hat{k}^2 + \frac{1}{12} \hat{k}^{(4)} \right) \left[ \hat{k}^2 + \frac{1}{12} \left( (\hat{k}^2)^2 + \hat{k}^{(4)} \right) \right. \\ &\quad \left. + \frac{1}{12} \left( (\hat{k}^2)^3 - \hat{k}^2 \hat{k}^{(4)} + 2\hat{k}^{(6)} \right) \right] + \frac{3}{12} \hat{k}^2 \prod_{\rho=0}^3 \hat{k}_{\rho}^2, \end{aligned} \quad (\text{B.11})$$

$$\begin{aligned}
A_{\mu\nu}^{(sq)}(k) = & \frac{1 - \delta_{\mu\nu}}{\Delta^{(sq)}} \left\{ (\tilde{k}^2 + \frac{1}{12}\tilde{k}^{(4)}) \left[ \tilde{k}^2 + \frac{1}{12}\tilde{k}^{(4)} \right] \right. \\
& + \frac{1}{12} \left( \tilde{k}^2 - \frac{1}{24}(\tilde{k}^2)^2 + \frac{1}{8}\tilde{k}^{(4)} \right) Q_{\mu\nu} + \frac{1}{12}\tilde{k}^2 Q_{\mu\nu}^2 - \frac{1}{48}Q_{\mu\nu}^3 \\
& + P_{\mu\nu} \left[ \frac{1}{144} \left( -(\tilde{k}^2)^2 - \frac{1}{8}\tilde{k}^2\tilde{k}^{(4)} - \frac{1}{144}(\tilde{k}^2)^2\tilde{k}^{(4)} + \frac{1}{12}(\tilde{k}^{(4)})^2 + \frac{1}{8}\tilde{k}^{(6)} \right) \right. \\
& + \frac{1}{24} \left( \tilde{k}^2 + \frac{1}{8}\tilde{k}^{(4)} + \frac{1}{288}\tilde{k}^2\tilde{k}^{(4)} \right) Q_{\mu\nu} + \frac{1}{864} \left( \tilde{k}^2 - \frac{1}{8}\tilde{k}^{(4)} \right) Q_{\mu\nu}^2 - \frac{1}{144}Q_{\mu\nu}^3 \\
& \left. + P_{\mu\nu}^2 \left( \frac{1}{864}(-\tilde{k}^2 + \frac{1}{8}\tilde{k}^{(4)}) + \frac{1}{144}Q_{\mu\nu} \right) \right\}, \quad (B.12)
\end{aligned}$$

$$\Delta^{(sq)} = \left( \tilde{k}^2 + \frac{1}{12}\tilde{k}^{(4)} \right)^2 \prod_{\rho=0}^3 \left( 1 + \frac{1}{12}\tilde{k}^2 \right). \quad (B.13)$$

In these formulas we used lattice units. The  $a$  dependence can be reinstated by substituting  $\tilde{k} \rightarrow a\tilde{k}$ .

For special momenta and polarizations the propagator greatly simplifies [27, 78]. From eqs. (B.1), (B.2) one deduces that vectors  $\varepsilon$  satisfying the following two conditions are eigenvectors of  $(D^{-1})_{\mu\nu}(k)$ .

1. The gauge condition  $\sum_{\mu} \tilde{k}_{\mu} \varepsilon_{\mu} = 0$  is satisfied;
2.  $\exists p \in \mathbb{R}$  (or  $\mathbb{C}$ ) such that  $\forall \mu \in \{\nu | \varepsilon_{\nu} \neq 0\}$   $k_{\mu} \in \{p, -p\}$ .

(Solutions to these conditions are for example  $\{\varepsilon = (1, 0, 0, 0)\}$  and  $\{k = (0, k_1, k_2, k_3)\}$  and  $\{\varepsilon = (0, 1, -1, 0)\}$ ;  $k = (k_0, k_1, k_1, k_3)\}$ ). Moreover, the eigenvalue equals simply  $\sum_{\rho} s_{\bar{\mu}\rho}(k) \tilde{k}_{\rho}^2$ , where  $\bar{\mu}$  is an arbitrary element of  $\{\mu | \varepsilon_{\mu} \neq 0\}$ . Therefore, whenever the two conditions are satisfied,  $\varepsilon$  is an eigenvector of the propagator with eigenvalue

$$\frac{\sum_{\mu\nu} \varepsilon_{\mu} D_{\mu\nu}(k) \varepsilon_{\nu}}{\sum_{\mu} \varepsilon_{\mu}^2} = d_{\bar{\mu}}(k), \quad d_{\bar{\mu}}(k) \equiv \frac{1}{\sum_{\rho} s_{\bar{\mu}\rho}(k) \tilde{k}_{\rho}^2}. \quad (B.14)$$

In section 5.5 we need the (tree-level) energy  $E_0(\mathbf{k})$  for  $d_1(k)$ , as well as the associated wave function renormalization  $Z_0(\mathbf{k})$ . They are defined through (cf. subsection 5.5.2)

$$d_1(k) = \frac{Z_0(\mathbf{k})}{k_0^2 + E_0^2(\mathbf{k})} + (\text{regular in } k_0), \quad (B.15)$$

for  $k$  close to the physical mass shell. An equivalent definition of  $Z_0(\mathbf{k})$  is

$$Z_0^{-1}(\mathbf{k}) = \frac{1}{2k_0} \frac{\partial}{\partial k_0} d_1^{-1}(k) \Big|_{k_0 = \pm i E_0(\mathbf{k})}. \quad (B.16)$$

Analytic expressions for  $E_0(\mathbf{k})$  and  $Z_0(\mathbf{k})$  can be easily obtained because  $d_1^{-1}(k)$  is quadratic in  $k_0^2$  for any values of  $c_0, c_1, c_2$  and  $c_4$ . When  $c_0$  is fixed by eq. (5.2.3) the results read

$$\begin{aligned}
E_0(\mathbf{k}) &= 2 \operatorname{asinh} \frac{1}{2} \sqrt{-\frac{B}{2A} \left( 1 - \sqrt{1 + 4AC/B^2} \right)}, \\
Z_0(\mathbf{k}) &= \left( \frac{\sinh E_0(\mathbf{k})}{E_0(\mathbf{k})} B \sqrt{1 + 4AC/B^2} \right)^{-1}, \quad (B.17)
\end{aligned}$$

where ( $\bar{c}_1 \equiv c_1 + 4c_4$ )

$$\begin{aligned} A &= \bar{c}_1 - c_4 \hat{k}_1^2, \\ B &= 1 - (\bar{c}_1 - c_2) \hat{k}_1^2 - 2c_2 \hat{k}^2, \\ C &= (1 - c_2 \hat{k}^2) \hat{k}^2 - (\bar{c}_1 - c_2) (\hat{k}_1^2 \hat{k}^2 + \hat{k}^{(4)}) + c_4 \hat{k}_1^2 \hat{k}^{(4)}, \end{aligned} \quad (\text{B.18})$$

with  $\hat{k}^{(4)} = \sum_{j=1}^3 k_j^4$ . For  $c_4 = 0$ ,  $E_0(\mathbf{k})$  reduces to the expression found in ref. [27], where however  $Z_0(\mathbf{k})$  is incorrect as it has the wrong limit for  $k \rightarrow 0$  (which corresponds to the continuum limit). The small- $k$  behavior of  $E_0(\mathbf{k})$  and  $Z_0(\mathbf{k})$  is

$$\begin{aligned} E_0(\mathbf{k}) &= |\mathbf{k}| \left[ 1 - \frac{1}{2}(\bar{c}_1 - c_2 + \frac{1}{12}) \left( \mathbf{k}^2 + \frac{\mathbf{k}^{(4)}}{\mathbf{k}^2} \right) + \mathcal{O}(k^4) \right], \\ Z_0(\mathbf{k}) &= 1 - 2(\bar{c}_1 - c_2 + \frac{1}{12}) \mathbf{k}^2 + (\bar{c}_1 - c_2) k_1^2 + \mathcal{O}(k^4). \end{aligned} \quad (\text{B.19})$$

Note that unlike in the continuum,  $Z_0(\mathbf{k})$  need not be 1.

## Appendix C: Vertex components

In this appendix we list the vertex components defined in eqs. (A.15)–(A.18). We restrict ourselves to the contributions of Wilson loops with non-zero tree-level coefficients in the square action, i.e. the  $1 \times 1$  ( $i = 0$ ),  $1 \times 2$  ( $i = 1$ ) and  $2 \times 2$  ( $i = 4$ ) loops. The results for  $i = 0$  and  $i = 1$ , as well as for the non-planar cases  $i = 2$  and  $i = 3$  were obtained before by Weisz and Wohlert [78]. We have greatly benefited from comparing to their expressions. For our calculation we used Mathematica [61]. We automated the translation to L<sup>A</sup>T<sub>E</sub>X, only doing by hand line-breaking and some minor changes to improve readability. We therefore expect the expressions below to be free of typesetting errors. We are also confident that no other mistakes were made because 1) the agreement with Weisz and Wohlert's expressions<sup>16</sup>; 2) the correctness of our expressions for small lattice spacings  $a$ . An example of the second point is given at the end of this appendix.

For shortness we suppress the momentum arguments ( $k_1, \dots, k_n$ ) on the left hand sides of the formulas below, and set  $a = 1$ . The lattice spacing can be reinstated by dimensional analysis: the mass dimension of  $k_i$  equals +1, that of any  $n$ -vertex component equals  $4 - n$ .

Due to the absence of Lorentz symmetry, explicit expressions for the vertices tend to be long, especially for  $i \neq 0$  and  $n \geq 4$ . We have put some effort in the search for relatively short forms. Nevertheless, we apologize for the lengthy formulas below.

<sup>16</sup>Apart from some more or less obvious typographic errors in ref. [78], we only disagree on the overall sign of the 3-vertex.

$$f_{\mu}^{(3,i=0)} = 0, \quad (\text{C.1})$$

$$f_{\mu}^{(3,i=1)} = -i \left\{ \hat{k}_1^2 c_{1\mu} (\widehat{k_2 - k_3})_{\mu} + 2 \text{ cyclic perms of the momenta} \right\}, \quad (\text{C.2})$$

$$f_{\mu}^{(3,i=4)} = -i \left\{ (2\hat{k}_1)^2 c_{1\mu} (\widehat{k_2 - k_3})_{\mu} + 2 \text{ cyclic perms of the momenta} \right\}, \quad (\text{C.3})$$

$$g_{\mu\nu}^{(3,i=0)} = -i c_{3\mu} (\widehat{k_1 - k_2})_{\nu}, \quad (\text{C.4})$$

$$g_{\mu\nu}^{(3,i=1)} = -i \left\{ 4 \left( \cos(k_{3\mu}) c_{1\mu} c_{2\mu} + \cos \frac{1}{2}(k_1 - k_2)_{\nu} c_{3\mu} c_{3\nu} \right) (\widehat{k_1 - k_2})_{\nu} \right. \\ \left. - \frac{1}{2} (2\hat{k})_{3\mu} \hat{k}_{3\nu} (\widehat{k_1 - k_2})_{\mu} \right\}, \quad (\text{C.5})$$

$$g_{\mu\nu}^{(3,i=4)} = -i c_{3\nu} \left\{ 8 \cos(k_{3\mu}) c_{1\mu} c_{2\mu} (\widehat{2(k_1 - k_2)})_{\nu} \right. \\ \left. - (2\hat{k})_{3\mu} (2\hat{k})_{3\nu} (\widehat{k_1 - k_2})_{\mu} \right\}, \quad (\text{C.6})$$

$$f_{\mu}^{(4,i=0)} = \sum_{\rho=0}^3 \frac{1}{8} \left\{ (\widehat{k_1 + k_3})_{\rho}^2 - \frac{1}{2} (\widehat{k_1 + k_2})_{\rho}^2 - \frac{1}{2} (\widehat{k_1 + k_4})_{\rho}^2 + \hat{k}_{1\rho} \hat{k}_{2\rho} \hat{k}_{3\rho} \hat{k}_{4\rho} \right\}, \quad (\text{C.7})$$

$$f_{\mu}^{(4,i=1)} = \sum_{\rho=0}^3 \left\{ \left[ \frac{1}{12} \hat{k}_{1\rho}^2 \left( 40 - 2 \hat{k}_{1\mu}^2 - 3 \hat{k}_{2\mu}^2 - 3 \hat{k}_{4\mu}^2 - 3 (\widehat{k_1 + k_2})_{\mu}^2 \right. \right. \right. \\ \left. \left. - 3 (\widehat{k_1 + k_4})_{\mu}^2 - 2 \hat{k}_{1\rho}^2 \right) + \frac{1}{8} (\widehat{k_1 + k_2})_{\rho}^2 \left( -20 + 2 \hat{k}_{1\mu}^2 + 2 \hat{k}_{2\mu}^2 \right. \right. \\ \left. \left. + (\widehat{k_1 + k_2})_{\mu}^2 + 2 (\widehat{k_1 + k_4})_{\mu}^2 + (\widehat{k_1 + k_2})_{\rho}^2 \right) \right] \\ \left. + 3 \text{ cyclic perms of the momenta} \right\}, \quad (\text{C.8})$$

$$f_{\mu}^{(4,i=4)} = \sum_{\rho=0}^3 \left\{ \left[ \frac{1}{12} (2\hat{k})_{1\rho}^2 \left( 32 - 2 \hat{k}_{1\mu}^2 - 3 \hat{k}_{2\mu}^2 - 3 \hat{k}_{4\mu}^2 - 3 (\widehat{k_1 + k_2})_{\mu}^2 \right. \right. \right. \\ \left. \left. - 3 (\widehat{k_1 + k_4})_{\mu}^2 \right) + \frac{1}{8} (2\hat{k})_{\rho}^2 \left( -16 + 2 \hat{k}_{1\mu}^2 + 2 \hat{k}_{2\mu}^2 \right. \right. \\ \left. \left. + (\widehat{k_1 + k_2})_{\mu}^2 + 2 (\widehat{k_1 + k_4})_{\mu}^2 \right) \right] + 3 \text{ cyclic perms of momenta} \right\}, \quad (\text{C.9})$$

$$g_{\mu\nu}^{(4,i=0)} = \frac{1}{8} \left( c_{3\nu} (\widehat{k_1 - k_2})_{\nu} - c_{1\nu} (\widehat{k_2 - k_3})_{\nu} - \hat{k}_{1\nu} \hat{k}_{2\nu} \hat{k}_{3\nu} \right) \hat{k}_{4\mu}, \quad (\text{C.10})$$

$$g_{\mu\nu}^{(4,i=1)} = -c_{4\nu} \left( \cos(k_1 - k_3)_{\nu} (2\hat{k})_{2\nu} + \frac{1}{2} (2\hat{k})_{4\nu} \right) \hat{k}_{4\mu} + \cos \frac{1}{2}(k_1 - k_3)_{\nu} \hat{k}_{2\nu} \\ \times \left( \cos(k_{4\mu}) \left( 2 c_{1\mu} c_{3\mu} \hat{k}_{2\mu} - c_{2\mu} (\widehat{k_1 + k_3})_{\mu} \right) - 4 c_{1\mu} c_{2\mu} c_{3\mu} (2\hat{k})_{4\mu} \right) \\ - \frac{1}{12} (2\hat{k})_{4\mu} \hat{k}_{4\nu} \left( 16 c_{1\mu} c_{2\mu} c_{3\mu} - (\widehat{k_1 + k_3})_{\mu} \hat{k}_{2\mu} + 2 c_{2\mu} \hat{k}_{1\mu} \hat{k}_{3\mu} \right) \\ - \cos(k_{4\mu}) c_{2\mu} c_{2\nu} (\widehat{k_1 - k_3})_{\mu} (\widehat{k_1 - k_3})_{\nu}, \quad (\text{C.11})$$

$$\begin{aligned} \bar{g}_{\mu\nu}^{(4,i=4)} &= 2 \cos(k_1 - k_3)_\nu c_{4\nu} \widehat{(2k)}_{2\nu} \\ &\quad \times \left( \cos(k_{4\mu}) (2 c_{1\mu} c_{3\mu} \widehat{k}_{2\mu} - c_{2\mu} \widehat{(k_1 + k_3)}_\mu) - 4 c_{1\mu} c_{2\mu} c_{3\mu} \widehat{(2k)}_{4\mu} \right) \\ &\quad - \frac{1}{2} c_{4\nu} \widehat{(2k)}_{4\mu} \widehat{(2k)}_{4\nu} \left( 16 c_{1\mu} c_{2\mu} c_{3\mu} - \widehat{(k_1 + k_3)}_\mu \widehat{k}_{2\mu} + 2 c_{2\mu} \widehat{k}_{1\mu} \widehat{k}_{3\mu} \right) \\ &\quad - 2 \cos(k_{2\nu}) \cos(k_{4\mu}) c_{2\mu} c_{4\nu} \widehat{(k_1 - k_3)}_\mu \widehat{(2(k_1 - k_3))}_\nu, \end{aligned} \quad (C.12)$$

$$\bar{h}_{\mu\nu}^{(4,i=0)} = 2 \cos \frac{1}{2}(k_2 - k_4)_\mu \cos \frac{1}{2}(k_1 - k_3)_\nu, \quad (C.13)$$

$$\begin{aligned} \bar{h}_{\mu\nu}^{(4,i=1)} &= 8 c_{1\mu} c_{3\mu} \cos(k_2 - k_4)_\mu \cos \frac{1}{2}(k_1 - k_3)_\nu \\ &\quad + 8 c_{2\nu} c_{4\nu} \cos \frac{1}{2}(k_2 - k_4)_\mu \cos(k_1 - k_3)_\nu, \end{aligned} \quad (C.14)$$

$$\bar{h}_{\mu\nu}^{(4,i=4)} = 32 c_{1\mu} c_{3\mu} c_{2\nu} c_{4\nu} \cos(k_2 - k_4)_\mu \cos(k_1 - k_3)_\nu, \quad (C.15)$$

$$\bar{i}_{\mu\nu}^{(4,i=0)} = -\cos \frac{1}{2}(k_3 - k_4)_\mu \cos \frac{1}{2}(k_1 - k_2)_\nu + \frac{1}{2} \widehat{k}_{3\mu} \widehat{k}_{4\mu} \widehat{k}_{1\nu} \widehat{k}_{2\nu}, \quad (C.16)$$

$$\begin{aligned} \bar{i}_{\mu\nu}^{(4,i=1)} &= \left[ -4 c_{1\mu} c_{2\mu} \cos(k_3 - k_4)_\mu \cos \frac{1}{2}(k_1 - k_2)_\nu \right. \\ &\quad \left. - \frac{1}{2} \cos \frac{1}{2}(k_1 + k_2)_\nu \widehat{(k_1 - k_2)}_\mu \widehat{(2(k_3 - k_4))}_\mu \right. \\ &\quad \left. + c_{1\mu} c_{2\mu} \widehat{(2k)}_{3\mu} \widehat{(2k)}_{4\mu} \widehat{k}_{1\nu} \widehat{k}_{2\nu} \right] + (k_1 \leftrightarrow k_3, k_2 \leftrightarrow k_4, \mu \leftrightarrow \nu), \end{aligned} \quad (C.17)$$

$$\begin{aligned} \bar{i}_{\mu\nu}^{(4,i=4)} &= 4 \left( -4 \cos(k_3 - k_4)_\mu \cos(k_1 - k_2)_\nu + \widehat{(2k)}_{3\mu} \widehat{(2k)}_{4\mu} \widehat{(2k)}_{1\nu} \widehat{(2k)}_{2\nu} \right) \\ &\quad \times c_{1\mu} c_{2\mu} c_{3\nu} c_{4\nu} - 2 \left( c_{3\nu} c_{4\nu} \cos(k_1 + k_2)_\nu \widehat{(k_1 - k_2)}_\mu \widehat{(2(k_3 - k_4))}_\mu \right. \\ &\quad \left. + c_{1\mu} c_{2\mu} \cos(k_3 + k_4)_\mu \widehat{(k_3 - k_4)}_\nu \widehat{(2(k_1 - k_2))}_\nu \right) \\ &\quad - \frac{1}{2} \widehat{(k_1 - k_2)}_\mu \widehat{(2(k_3 + k_4))}_\mu \widehat{(k_3 - k_4)}_\nu \widehat{(2(k_1 + k_2))}_\nu, \end{aligned} \quad (C.18)$$

$$f_{\mu}^{(W,i)} = \sum_{\rho=0}^3 \widehat{k}_{1\rho} \widehat{k}_{2\rho} \widehat{k}_{3\rho} \widehat{k}_{4\rho} K_{\rho\mu}^{(i)}(k_1, k_2, k_3, k_4), \quad (C.19)$$

$$g_{\mu\nu}^{(W,i)} = -\widehat{k}_{4\mu} \widehat{k}_{1\nu} \widehat{k}_{2\nu} \widehat{k}_{3\nu} K_{\mu\nu}^{(i)}(k_1, k_2, k_3, k_4), \quad (C.20)$$

$$h_{\mu\nu}^{(W,i)} = \widehat{k}_{3\mu} \widehat{k}_{4\mu} \widehat{k}_{1\nu} \widehat{k}_{2\nu} K_{\mu\nu}^{(i)}(k_1, k_2, k_3, k_4). \quad (C.21)$$

In the last three equations we introduced

$$K_{\mu\nu}^{(i=0)}(k_1, k_2, k_3, k_4) \equiv 2, \quad (C.22)$$

$$K_{\mu\nu}^{(i=1)}(k_1, k_2, k_3, k_4) \equiv 32 (c_{1\mu} c_{2\mu} c_{3\mu} c_{4\mu} + c_{1\nu} c_{2\nu} c_{3\nu} c_{4\nu}), \quad (C.23)$$

$$K_{\mu\nu}^{(i=4)}(k_1, k_2, k_3, k_4) \equiv 512 (c_{1\mu} c_{2\mu} c_{3\mu} c_{4\mu} c_{1\nu} c_{2\nu} c_{3\nu} c_{4\nu}). \quad (C.24)$$

The 5-vertex component  $f^{(5,i)}$  is not needed in the calculation of the one-loop Symanzik coefficients. The other components are expressed below as polynomials in a small number of functions, namely  $\widehat{k}_{j\lambda}$ ,  $c_{j\lambda}$  and

$$\begin{aligned} c_{ij\lambda}^{\pm} &\equiv \cos \frac{1}{2}(k_i \pm k_j)_\lambda; \quad s_{ij\lambda}^{\pm} \equiv \frac{1}{2} \widehat{(k_i \pm k_j)}_\lambda, \\ c_{ij\lambda}^{2\pm} &\equiv \cos(k_i \pm k_j)_\lambda; \quad s_{ij\lambda}^{2\pm} \equiv \frac{1}{2} \widehat{(2(k_i \pm k_j))}_\lambda. \end{aligned} \quad (C.25)$$

For each component only those functions are used that are even or odd with respect to the symmetry operation belonging to that component (see eq. (A.19)).

$$g_{\mu\nu}^{(5,i=0)} = 4ic_{5\mu} \left\{ (-3c_{23\nu}^- + 4c_{23\nu}^+) s_{14\nu}^- - 3c_{14\nu}^- s_{23\nu}^- \right\}, \quad (C.26)$$

$$\begin{aligned} g_{\mu\nu}^{(5,i=1)} = & 8i \left\{ c_{5\mu} c_{5\nu} \left( (-3c_{23\nu}^- + 4c_{23\nu}^+) s_{14\nu}^2 - 3c_{14\nu}^- s_{23\nu}^2 \right) \right. \\ & + \cos(k_{5\mu}) \left( 3 \left( c_{23\nu}^- s_{14\nu}^- + c_{14\nu}^- s_{23\nu}^- \right) \left( -c_{23\mu}^- (c_{14\mu}^- + 2c_{14\mu}^+) \right) \right. \\ & - (2c_{14\mu}^- + 3c_{14\mu}^+) c_{23\mu}^+ + s_{14\mu}^- s_{23\mu}^- - s_{14\mu}^+ s_{23\mu}^+ \left. \right) \\ & + 2c_{23\mu}^+ s_{14\mu}^- \left( (c_{14\mu}^- + c_{14\mu}^+) (3c_{23\mu}^- + 5c_{23\mu}^+) - 3s_{14\mu}^- s_{23\mu}^- + s_{14\mu}^+ s_{23\mu}^+ \right) \\ & + 2c_{14\mu}^- \left( -3s_{23\mu}^- s_{14\mu}^+ + s_{14\mu}^- s_{23\mu}^+ \right) s_{23\nu}^+ \left. \right) \\ & + \frac{1}{2}(\widehat{2k})_{5\mu} \left( 6(-c_{23\nu}^- + c_{23\nu}^+) s_{14\nu}^- \left( -(c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^+ \right) \right. \\ & + (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^+ \left. \right) + \left( (3c_{23\mu}^- + 4c_{23\mu}^+) s_{14\mu}^- + 3c_{14\mu}^- s_{23\mu}^- \right) \\ & \times \left( c_{23\nu}^+ s_{14\nu}^+ + c_{14\nu}^+ s_{23\nu}^+ \right) + 6c_{14\nu}^- \left[ s_{23\nu}^- \left( (c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^+ \right) \right. \\ & - (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^+ \left. \right) \\ & \left. - \left( (c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^- + (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^- \right) s_{23\nu}^+ \right] \left. \right\}, \quad (C.27) \end{aligned}$$

$$\begin{aligned} g_{\mu\nu}^{(5,i=4)} = & 16ic_{5\nu} \left\{ \cos(k_{5\mu}) \left( 2c_{14\nu}^2 s_{23\nu}^+ - 3s_{23\mu}^- s_{14\mu}^+ + s_{14\mu}^- s_{23\mu}^+ \right) \right. \\ & + 3 \left( c_{23\nu}^2 s_{14\nu}^- + c_{14\nu}^2 s_{23\nu}^- \right) \left( -c_{23\mu}^- (c_{14\mu}^- + 2c_{14\mu}^+) \right) \\ & - (2c_{14\mu}^- + 3c_{14\mu}^+) c_{23\mu}^+ + s_{14\mu}^- s_{23\mu}^- - s_{14\mu}^+ s_{23\mu}^+ \left. \right) \\ & + 2c_{23\mu}^+ s_{14\mu}^- \left( (c_{14\mu}^- + c_{14\mu}^+) (3c_{23\mu}^- + 5c_{23\mu}^+) - 3s_{14\mu}^- s_{23\mu}^- + s_{14\mu}^+ s_{23\mu}^+ \right) \left. \right) \\ & + (\widehat{2k})_{5\mu} \left( 6c_{14\nu}^- \left( -c_{23\nu}^- + c_{23\nu}^+ \right) s_{14\nu}^- \left( -(c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^+ \right) \right. \\ & + (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^+ \left. \right) + \left( (3c_{23\mu}^- + 4c_{23\mu}^+) s_{14\mu}^- + 3c_{14\mu}^- s_{23\mu}^- \right) \\ & \times \left( c_{23\nu}^+ s_{14\nu}^+ + c_{14\nu}^+ s_{23\nu}^+ \right) + 6c_{14\nu}^- \left[ c_{23\nu}^- s_{23\mu}^- \left( (c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^+ \right) \right. \\ & - (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^+ \left. \right) - c_{23\mu}^- \left( (c_{23\mu}^- + c_{23\mu}^+) s_{14\mu}^- \right) \\ & \left. + (c_{14\mu}^- + c_{14\mu}^+) s_{23\mu}^- \right] s_{23\nu}^+ \left. \right\}, \quad (C.28) \end{aligned}$$

$$h_{\mu\nu}^{(5,i=0)} = -24is_{24\mu}^- \left( c_{3\nu} c_{15\nu}^+ + \frac{1}{2} \widehat{k}_{3\nu} s_{15\nu}^+ \right), \quad (C.29)$$

$$\begin{aligned} h_{\mu\nu}^{(5,i=1)} = & -48i \left\{ (c_{24\nu}^- + c_{24\nu}^+) s_{24\mu}^- \left( \cos(k_{3\nu}) c_{15\mu}^2 + \frac{1}{2}(\widehat{2k})_{3\nu} s_{15\mu}^2 \right) \right. \\ & + 2c_{3\mu}^- (c_{15\mu}^- + c_{15\mu}^+) s_{24\mu}^- \left( c_{3\nu} c_{15\nu}^+ + \frac{1}{2} \widehat{k}_{3\nu} s_{15\nu}^+ \right) \\ & \left. + 2c_{24\mu}^- c_{3\mu} s_{15\mu}^- \left( c_{3\nu} c_{15\nu}^+ + \frac{1}{2} \widehat{k}_{3\nu} s_{15\nu}^+ \right) \right\}, \quad (C.30) \end{aligned}$$

$$h_{\mu\nu}^{(5,i=4)} = -192i c_{3\nu} (c_{24\nu}^- + c_{24\nu}^+) (s_{24\mu}^2 (c_{15\mu}^- + c_{15\mu}^+) + c_{24\mu}^2 s_{15\mu}^-) \\ \times (\cos(k_{3\nu}) c_{15\mu}^+ + \frac{1}{2} (2\hat{k})_{3\nu} s_{15\mu}^2), \quad (C.31)$$

$$h_{\mu\nu}^{(5,i=0)} = 8i (c_{45\nu}^+ s_{45\mu}^- + \frac{1}{2} \hat{k}_{2\nu} s_{13\nu}^- s_{45\mu}^+), \quad (C.32)$$

$$h_{\mu\nu}^{(5,i=1)} = 16i \left\{ (3c_{13\mu}^- c_{2\mu} + c_{45\mu}^+) c_{45\nu}^+ s_{45\mu}^2 + 3c_{2\mu} (-c_{13\nu}^- c_{2\nu} c_{45\mu}^2 + (-c_{45\mu}^- + c_{45\mu}^+) c_{45\nu}^+) s_{13\mu}^- + c_{45\nu}^2 (c_{45\mu}^- + c_{45\mu}^+) s_{45\mu}^- + \frac{1}{2} \hat{k}_{2\nu} s_{13\nu}^- \right. \\ \left. \times [(c_{13\mu}^- + c_{13\mu}^+) (2c_{2\mu} s_{45\mu}^2 + \frac{1}{2} c_{45\mu}^2 \hat{k}_{2\mu}) + c_{2\mu} c_{45\mu}^2 s_{13\mu}^-] + \frac{3}{2} (c_{45\mu}^- + c_{45\mu}^+) \right. \\ \left. \times (2\hat{k})_{2\nu} s_{13\nu}^- s_{45\mu}^+ - s_{45\nu}^- (\frac{3}{2} c_{13\nu}^- (2\hat{k})_{2\nu} + s_{46\nu}^2) s_{45\mu}^+ \right\}, \quad (C.33)$$

$$h_{\mu\nu}^{(5,i=4)} = 32i \left\{ c_{45\nu}^2 (3c_{13\mu}^- c_{2\mu} + c_{45\mu}^+) (c_{45\nu}^- + c_{45\nu}^+) s_{45\mu}^2 - 3c_{2\mu} (c_{13\nu}^- \cos(k_{2\nu}) c_{45\mu}^2 + (c_{45\mu}^- - c_{45\mu}^+) c_{45\nu}^2) (c_{45\nu}^- + c_{45\nu}^+) s_{13\mu}^- \right. \\ \left. + \frac{3}{2} (c_{45\nu}^- + c_{45\nu}^+) (2\hat{k})_{2\nu} s_{13\nu}^- [(c_{13\mu}^- + c_{13\mu}^+) (2c_{2\mu} s_{45\mu}^2 - \frac{1}{2} c_{45\mu}^2 \hat{k}_{2\mu}) \right. \\ \left. + c_{2\mu} c_{45\mu}^2 s_{13\mu}^-] - s_{45\nu}^- (3\cos(k_{2\nu}) c_{2\mu} c_{45\mu}^2 s_{13\nu}^- s_{13\mu}^- + (3c_{13\mu}^- c_{2\mu} + c_{45\mu}^+) s_{45\mu}^2 s_{45\nu}^2 + \frac{3}{2} c_{13\nu}^- (2\hat{k})_{2\nu} \right. \\ \left. \times [(c_{13\mu}^- + c_{13\mu}^+) (2c_{2\mu} s_{45\mu}^2 - \frac{1}{2} c_{45\mu}^2 \hat{k}_{2\mu}) + c_{2\mu} c_{45\mu}^2 s_{13\mu}^-] \right\}. \quad (C.34)$$

A powerful tool for tracking errors is the check whether the  $a \rightarrow 0$  behavior of the vertices is consistent with eq. (5.2.1). As a non-trivial example we perform this analysis for the 5-vertex.

Reinstating the lattice spacing  $a$ , the  $c_i$ -weighted sums of eqs (C.26)–(C.34) equal, up to  $\mathcal{O}(a^4)$  corrections,

$$g_{\mu\nu}^{(5)} = 2ia^2 (c_0 + 20c_1 + 64c_4) \{ (k_1 - k_4)_\nu - 3(k_2 - k_3)_\nu \}, \\ h_{\mu\nu}^{(5)} = -12ia^2 \{ (c_0 + 20c_1 + 64c_4) (k_2 - k_4)_\mu + 4(c_1 + 4c_4) (k_1 - k_5)_\mu \}, \\ h_{\mu\nu}'^{(5)} = 4ia^2 \{ (c_0 + 20c_1 + 64c_4) (k_4 - k_5)_\mu - 6(c_1 + 4c_4) (k_1 - k_3)_\mu \}. \quad (C.35)$$

Eq. (5.2.1) seems to be violated because  $h$  and  $h'$  are not proportional to  $(c_0 + 20c_1 + 64c_4)$ . However, we should take into account the remark below eq. (5.3.2): eq. (5.2.1) only holds in terms of the field  $\bar{A}_\mu$ , defined by

$$U_\mu(x) = P \exp \left( \int_0^a ds \bar{A}_\mu(x + s\hat{\mu}) \right) \\ = 1 + \sum_{m=1}^{\infty} \int_{0 < s_1 < s_2 < \dots < s_m < a} d^m s \bar{A}_\mu(x + s_1\hat{\mu}) \cdots \bar{A}_\mu(x + s_m\hat{\mu}). \quad (C.36)$$

By expanding eqs. (C.36), (5.1.1) with respect to  $a$ , the relation between  $\bar{A}_\mu$  and  $A_\mu$  is found to be

$$A_\mu(x) = \bar{A}_\mu(x') + \frac{1}{12} a^2 \left\{ \frac{1}{2} \partial_\mu^2 \bar{A}_\mu(x') + [\bar{A}_\mu(x'), \partial_\mu \bar{A}_\mu(x')] \right\} + \mathcal{O}(a^3), \quad (C.37)$$

where  $x' = x + \frac{1}{2}a\bar{\mu}$ . Below we neglect the shift over  $\frac{1}{2}a\bar{\mu}$ , because it only contributes total derivatives to the Lagrangian density that hence drop out of the action.

The commutator term in eq. (C.37) is interesting because it mixes  $n$  and  $n+1$  vertices. In particular if we expand eq. (A.6) with respect to  $\bar{A}$  instead of  $A$ , denoting terms by  $\bar{S}_n$ , we find the relation

$$\frac{g_0^3}{5!} \bar{S}_5 = \frac{g_0^3}{5!} \bar{S}_5|_{A_\mu \rightarrow \bar{A}_\mu} + \frac{g_0^2}{4!} \left( S_4|_{A_\mu \rightarrow \bar{A}_\mu + \frac{1}{2}a^2[\bar{A}_\mu, \partial_\mu \bar{A}_\mu]} \right)_{\text{terms} \sim \bar{A}^5} + \mathcal{O}(a^4). \quad (\text{C.38})$$

It is clear that only the  $S_4$  components with a non-vanishing continuum limit, i.e.  $h^{(4)}$  and  $\bar{h}^{(4)}$ , contribute to  $\bar{S}_5$  at order  $a^2$ . By comparing the Lorentz-index structures of eqs. (A.16) and (A.18) one easily deduces that only  $\bar{h}^{(5)}$  and  $\bar{h}'^{(5)}$  are affected by this contribution, where of course  $\bar{h}^{(5)}$ ,  $\bar{h}'^{(5)}$ , ... have the same meaning with respect to  $\bar{A}$  as  $h^{(5)}$ ,  $h'^{(5)}$ , ... have with respect to  $A$ . After some algebra, including the disentanglement of the  $SU(N)$  structure, one finds the explicit result

$$\begin{aligned} \bar{g}_{\mu\nu}^{(5)} &= 2ia^2(c_0 + 20c_1 + 64c_4) \{ (k_1 - k_4)_\nu - 3(k_2 - k_3)_\nu \} + \mathcal{O}(a^4), \\ \bar{h}_{\mu\nu}^{(5)} &= -4ia^2(c_0 + 20c_1 + 64c_4) \{ 3(k_2 - k_4)_\mu + (k_1 - k_5)_\mu \} + \mathcal{O}(a^4), \\ \bar{h}'_{\mu\nu}{}^{(5)} &= 2ia^2(c_0 + 20c_1 + 64c_4) \{ 2(k_4 - k_5)_\mu - (k_1 - k_3)_\mu \} + \mathcal{O}(a^4), \end{aligned} \quad (\text{C.39})$$

which is consistent with eq. (5.2.1) (as it should be). We stress that this is a rather stringent test on the structure of  $g^{(5,i)}$ ,  $h^{(5,i)}$  and  $h'^{(5,i)}$ , and in particular on the signs and numerical prefactors of  $h^{(5,i)}$  and  $h'^{(5,i)}$ .

## Appendix D: Feynman rules with a twist

In a periodic finite volume the Fourier representations given in appendix A are still valid, as long as one replaces the integral (A.2) by a sum. In a twisted finite volume the situation is different due to color-momentum mixing. The Fourier representation appropriate to the 'twisted tube' (see section 5.5) is given below.

We use lattice units in this appendix. For shortness we adopt the summation convention, both for Lorentz indices and for 'color momentum'  $e$  in the twisted directions 1 and 2 ( $e_\nu/m$  ( $\nu = 1, 2$ ) runs over  $\{0, 1, \dots, N-1\}$ , where  $m \equiv 2\pi/(NL)$ ).

$$S_{\text{measure}} = \frac{N}{24} g_0^2 \delta_{\mu_1 \mu_2} \sum_{k_1, k_2} \delta(k_1 + k_2) \bar{A}_{\mu_1}(k_1) \bar{A}_{\mu_2}(k_2) (-2z^{\frac{1}{2}(k, k)}) + \mathcal{O}(g_0^4), \quad (\text{D.1})$$

$$\begin{aligned} S_{\text{ghost}} &= \sum_{k_1, k_2} \bar{c}(k_1) \bar{c}(k_2) \left[ \delta(k_1 + k_2) \bar{k}_1^2 \right. \\ &\quad + i g_0 \sum_{k_3} \delta(k_1 + k_2 + k_3) \bar{A}_\mu(k_3) f(k_1, k_2, k_3) z^{-\frac{1}{2}(k_1, k_1)} \bar{k}_{1\mu} c_{2\mu} \\ &\quad \left. + \frac{1}{12} g_0^2 \delta_{\mu_3 \mu_4} \sum_{k_3, k_4} \delta(k_1 + k_2 + k_3 + k_4) \bar{A}_{\mu_3}(k_3) \bar{A}_{\mu_4}(k_4) \right. \\ &\quad \left. \times f(k_1, k_3, e) f(k_2, k_4, -e) z^{-\frac{1}{2}(e, e)} z^{-\frac{1}{2}(k_1, k_1)} \bar{k}_{1\mu_3} \bar{k}_{2\mu_4} + \mathcal{O}(g_0^4) \right], \quad (\text{D.2}) \end{aligned}$$

$$S_2(\{c_i\}) = \oint_{k_1, k_2} \delta(k_1 + k_2) \bar{A}_{\mu_1}(k_1) \bar{A}_{\mu_2}(k_2) (-2z^{\frac{1}{2}\langle k, k \rangle}) (D^{-1})_{\mu_1 \mu_2}(k_1), \quad (D.3)$$

$$S_3(\{c_i\}) = \oint_{k_1, k_2, k_3} \delta(k_1 + k_2 + k_3) \bar{A}_{\mu_1}(k_1) \bar{A}_{\mu_2}(k_2) \bar{A}_{\mu_3}(k_3) \\ \times (-2f(k_1, k_2, k_3)) V_{\mu_1 \mu_2 \mu_3}(k_1, k_2, k_3), \quad (D.4)$$

$$S_4(\{c_i\}) = \oint_{k_1, k_2, k_3, k_4} \delta(k_1 + k_2 + k_3 + k_4) \bar{A}_{\mu_1}(k_1) \bar{A}_{\mu_2}(k_2) \bar{A}_{\mu_3}(k_3) \bar{A}_{\mu_4}(k_4) \\ \times [-2f(k_1, k_2, e) f(k_3, k_4, -e) z^{-\frac{1}{2}\langle e, e \rangle}] \\ \times (V_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4) - V_{\mu_2 \mu_1 \mu_3 \mu_4}(k_2, k_1, k_3, k_4)) \\ - S(k_1, k_2, k_3, k_4) W_{\mu_1 \mu_2 \mu_3 \mu_4}(k_1, k_2, k_3, k_4)], \quad (D.5)$$

$$S_5(\{c_i\}) = \oint_{k_1, k_2, k_3, k_4, k_5} \delta(k_1 + k_2 + k_3 + k_4 + k_5) \bar{A}_{\mu_1}(k_1) \cdots \bar{A}_{\mu_5}(k_5) \\ \times C(k_1, k_2, k_3, k_4, k_5) V_{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5}(k_1, k_2, k_3, k_4, k_5). \quad (D.6)$$

In these equations,  $D_{\mu\nu}$  and the vertices  $V$ ,  $W$  are precisely the same as in infinite volume. However, the color factors are now functions of the momenta (in the twisted directions only),

$$f(k_1, k_2, k_3) \equiv \frac{1}{N} \text{Tr}([\Gamma_{k_1}, \Gamma_{k_2}] \Gamma_{k_3}) \\ = 2i(1 - \chi_{k_1 + k_2 + k_3}) \\ \times z^{\frac{1}{2}\langle k_1, k_1 \rangle + \frac{1}{2}\langle k_2, k_2 \rangle + \frac{1}{2}\langle k_3, k_3 \rangle} \sin\left(\frac{\pi}{N} \langle k_1, k_2 \rangle\right), \quad (D.7)$$

$$S(k_1, k_2, k_3, k_4) \equiv \frac{1}{24N} \text{Tr}(\Gamma_{k_1} \Gamma_{k_2} \Gamma_{k_3} \Gamma_{k_4} + 23 \text{ permutations}) \\ = \frac{1}{3} (1 - \chi_{k_1 + k_2 + k_3 + k_4}) z^{-\frac{1}{2} \sum_{1 \leq i < j \leq 4} \langle k_i, k_j \rangle} \\ \times \left\{ z^{\frac{1}{2}\langle k_1 + k_2, k_3 + k_4 \rangle} \cos\left(\frac{\pi}{N} \langle k_1, k_2 \rangle\right) \cos\left(\frac{\pi}{N} \langle k_3, k_4 \rangle\right) \right. \\ \left. + (k_2 \leftrightarrow k_3) + (k_2 \leftrightarrow k_4) \right\}, \quad (D.8)$$

$$C(k_1, k_2, k_3, k_4, k_5) \equiv \frac{1}{N} \text{Tr}(\Gamma_{k_1} \Gamma_{k_2} \Gamma_{k_3} \Gamma_{k_4} \Gamma_{k_5} - \Gamma_{k_5} \Gamma_{k_4} \Gamma_{k_3} \Gamma_{k_2} \Gamma_{k_1}) \\ = 2i(1 - \chi_{k_1 + k_2 + k_3 + k_4 + k_5}) z^{-\frac{1}{2} \sum_{1 \leq i < j \leq 5} \langle k_i, k_j \rangle} \\ \times \sin\left(\frac{\pi}{N} \sum_{1 \leq i < j \leq 5} \langle k_i, k_j \rangle\right). \quad (D.9)$$

The notation is explained in subsection 5.5.1. Note that the factor  $(-2z^{\frac{1}{2}\langle k, k \rangle})$  appearing in  $S_2\{c_i\}$  is a color factor, brought about by  $N^{-1} \text{Tr}(\Gamma_k \Gamma_{-k}) = z^{\frac{1}{2}\langle k, k \rangle}$  as opposed to  $\text{Tr}(T^a T^b) = -\frac{1}{2} \delta_{ab}$ .

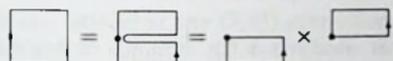
## Appendix E: Positivity of the square action

For many values of the coefficients  $c_i(g_0^2)$ , the action (4.2.1) is not positive [25]. Such a choice of coefficients is unacceptable because the 'vacuum'  $A_\mu = 0$  (i.e.  $U_\mu = 1$ ), which has zero action, would not be a minimal-action configuration, and hence not a correct field configuration to expand about in perturbation theory. It is therefore important to prove that the square action is positive.

The basic ingredient [25] for the proof is the statement that for any two unitary matrices  $U$  and  $V$

$$\operatorname{Re} \operatorname{Tr} (1 - UV) \leq 2\operatorname{Re} \operatorname{Tr} (1 - U) + 2\operatorname{Re} \operatorname{Tr} (1 - V). \quad (\text{E.1})$$

This can be directly applied to eq. (4.2.1) by expressing the  $1 \times 2$ ,  $2 \times 2$  and corner Wilson loops as products of smaller loops (we choose  $c_3 = 0$ ). For example, one may write



$$\square = \text{rectangle} = \text{horizontal loop} \times \text{vertical loop}. \quad (\text{E.2})$$

Noting that any unitary matrix  $U$  satisfies

$$\operatorname{Re} \operatorname{Tr} (1 - U) \geq 0, \quad (\text{E.3})$$

one quickly derives a rough lower bound:

$$S(\{c_i(g_0^2)\}) \geq \left( c_0(g_0^2) + 8(c_1(g_0^2) + 2c_4(g_0^2)) + \frac{80}{3}c_2(g_0^2) \right) S_{\text{Wilson}}, \quad (\text{E.4})$$

where  $\bar{x} \equiv \min(x, 0)$ . The Wilson action  $S_{\text{Wilson}}$  is positive due to eq. (E.3). Therefore the action (4.2.1) is positive if the prefactor on the right hand side of eq. (E.4) is.

At tree level, the Lüscher-Weisz [25] as well as the square action satisfy eq. (E.4) with a positive prefactor. If one includes the one-loop corrections that are computed in this chapter, table 5-4, the prefactor is still positive for any reasonable value of  $g_0^2$ , as long as the free parameter  $c_4$  is chosen not too big. For example, the one-loop square action is positive for  $c_4 = 0$ ;  $0 < g_0^2 < 10.9$ , and also for  $g_0^2 = 1$ ;  $144 c_4 < 7.2$  ( $N = 3$ ).



## 6 Conclusion

In this thesis we have considered various aspects of lattice artefacts. Also, we have investigated instantons and sphalerons. Here we summarize our results, mentioning some potential applications.

In chapter 2 we showed that lattice artefacts are not necessarily a nuisance. They can be put to something useful: the numerical search for large instantons in pure gauge theories. In a classical setting, we determined the  $\mathcal{O}(a^2)$  artefacts analytically, and tuned them by adding a  $2 \times 2$  Wilson loop to the Wilson action. In this way we obtained an 'over-improved' action, for which instantons grow upon cooling (i.e. systematically lowering the action), at least above a certain initial size.

We also presented the  $\mathcal{O}(a^4)$  corrections for two special actions, and a systematic method to compute  $a^{2n}$  corrections in general. This has been applied to the construction of a family of actions free of  $\mathcal{O}(a^4)$  lattice artefacts in ref. [63]. One of these actions was used to measure the topological susceptibility through cooling of configurations generated by Monte Carlo simulations of the Wilson action. In this way the instanton size distribution can be determined much more reliably than is possible if the Wilson action is used for the cooling.

In chapter 2, we used *over-improvement* in the strict sense. By this method we obtained the widest  $SU(2)$  instantons on  $T^3 \times \mathbb{R}$ , to a good numerical precision. It was established that these instantons go through sphalerons (or, at least, saddle points of the potential with only one unstable mode). One might envisage the use of the instantons (or sphalerons) for an extension of the finite-volume calculation in ref. [8], to include extra non-perturbative effects (as was done on  $S^3$  in refs. [15, 16]). However, for this the numerical nature of our solutions is a drawback. Nevertheless, they might be useful to guide the search for analytic solutions.

The cooling method did not permit us to completely determine the instanton moduli-space. Because some interesting aspects are involved, we considered the  $O(3)$   $\sigma$ -model on  $T^1 \times \mathbb{R}$  in chapter 3. Due to a similar classical vacuum structure, it may be viewed as a toy model for  $SU(2)$  pure gauge theory on  $T^3 \times \mathbb{R}$ . In the  $\sigma$ -model we were able to do a completely analytic calculation. We proved that the configuration where the widest instanton crosses the barrier, is a sphaleron. Furthermore, the instanton size  $\rho$  was found to depend on the asymptotic values of the instanton field at  $t \rightarrow \pm\infty$ , which must lie in the vacuum valley. This is different to the situation for  $SU(2)$  on  $T^3 \times \mathbb{R}$ . In that case, general arguments seem to imply that the scale parameter can vary independent of the vacuum valley parameters, as was discussed in section 2.2.

In chapters 4 and 5 we studied the improvement of pure lattice gauge theory beyond tree level. We introduced the square action, which is an alternative to the Symanzik improved action due to Lüscher and Weisz [26]. The structure of its tree-

level coefficients is such that the bare propagator can be easily diagonalized with respect to the Lorentz indices. This is an advantage for perturbative calculations, which could be useful in a more general setting than the background field context of chapter 4. There we computed the one-loop effective potential for fields in the classical vacuum valley, in SU(2) pure gauge theory on the spatial geometry  $T^3$ .

In chapter 5 we computed the one-loop improvement coefficients for the square action. This allowed us to do a perturbative test of tadpole improvement, namely to check its prediction of the one-loop coefficients  $c'_1$  from the tree-level coefficients  $c_1$ . Tadpole improvement was known to behave well under a similar test for the Lüscher-Weisz action, giving for SU(3) 76% and for SU(2) 80% of the  $g_0^2$  term in the relevant ratio  $(c_1 + c'_1 g_0^2)/(c_0 + c'_0 g_0^2)$ . For the square action we found 79% and 80%, respectively. Hence the tadpole method stands the test well for both actions, which is less likely to be a coincidence. Our results thus support the claim [29] that tadpole improvement brings about an efficient resummation of conventional lattice perturbation theory.

Nevertheless, in chapter 4 we found that tadpole improvement behaves less favorably in a non-perturbative context, at least in finite volumes. First, for the vacuum-valley effective potential we saw that it cannot give a substantial additional improvement over the (square) Symanzik action. Moreover, also our finite-volume Monte Carlo simulations showed that the tadpole action does not improve on the Symanzik action, even if the latter leaves ample room for improvement. Although tadpole improvement works well in many other cases (see e.g. ref. [31]), our results give a warning. As long as the inability of the tadpole method to improve certain quantities is not understood, its use in Monte Carlo simulations is dangerous.

Returning to the square action, we think it is a welcome cousin to the one-loop Lüscher-Weisz action. By using the two side by side, tests of Symanzik improvement, or of schemes built on Symanzik improvement, will become more convincing. An example is the above perturbative test of tadpole improvement.

## References

- [1] H.D. Politzer, Phys. Rev. Lett. 30 (1973) 1346; D.J. Gross and F. Wilczek, Phys. Rev. Lett. 30 (1973) 1343; S. Coleman and D.J. Gross, Phys. Rev. Lett. 31 (1973) 851.
- [2] J. Sexton et al., Phys. Rev. Lett. 75 (1995) 4563.
- [3] G. Bali et al., Phys. Lett. B309 (1993) 378.
- [4] G. 't Hooft, Nucl. Phys. B33 (1971) 173.
- [5] K. Wilson and J. Kogut, Phys. Rep. 12C (1974) 75; K. Wilson, Rev. Mod. Phys. 47 (1975) 773; K. Wilson, Rev. Mod. Phys. 55 (1983) 583.
- [6] T. Reisz, Nucl. Phys. B318 (1989) 417.
- [7] K. Wilson, Phys. Rev. D10 (1974) 2445.
- [8] M. Lüscher, Nucl. Phys. B219 (1983) 233; P. van Baal and J. Koller, Ann. Phys. (N.Y.) 174 (1987) 299; J. Koller and P. van Baal, Nucl. Phys. B302 (1988) 1.
- [9] N.M. Christ and T.D. Lee, Phys. Rev. D22 (1980) 939.
- [10] G. 't Hooft, Nucl. Phys. B153 (1979) 141; G. 't Hooft, Acta Physica Aust. Suppl. XXII (1980) 531; P. van Baal, Comm. Math. Phys. 85 (1982) 529.
- [11] P. van Baal, Phys. Lett. B224 (1989) 397; Nucl. Phys. B351 (1991) 183.
- [12] C. Michael et al., Phys. Lett. B207 (1988) 313; C. Michael, Nucl. Phys. B329 (1990) 225; B. Berg and A. Billoire, Phys. Rev. D40 (1989) 550.
- [13] N. Manton, Phys. Rev. D28 (1983) 2019, F. R. Klinkhamer and N. Manton, Phys. Rev. D30 (1984) 2212.
- [14] C.H. Taubes, Comm. Math. Phys. 86 (1982) 257, 299; C.H. Taubes, in: *Progress in gauge field theory*, eds. G. 't Hooft et al. (Plenum Press, New York, 1984) p. 563.
- [15] B. van den Heuvel, *Non-perturbative phenomena in gauge theory on  $S^3$*  (Ph.D. thesis, Leiden, September 1996).
- [16] P. van Baal and N.D. Hari Dass, Nucl. Phys. B385 (1992) 185.
- [17] A.A. Belavin et al., Phys. Lett. B59 (1975) 85.
- [18] G. 't Hooft, Phys. Rev. D14 (1976) 3432.
- [19] G. 't Hooft, Phys. Rev. Lett. 37 (1976) 8; R. Jackiw and C. Rebbi, Phys. Rev. Lett. 37 (1976) 172.
- [20] H.J. Rothe, *Lattice Gauge Theories* (World Scientific, Singapore, 1992); J. Smit, lecture notes (University of Amsterdam).
- [21] A.D. Sokal, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 55.
- [22] K. Wilson, in: *Recent developments in gauge theories*, eds. G. 't Hooft et al. (Plenum Press, New York, 1980).
- [23] P. Hasenfratz and F. Niedermayer, Nucl. Phys. B414 (1994) 785; T. DeGrand et al., Phys. Lett. B365 (1996) 233.
- [24] K. Symanzik, Nucl. Phys. B226 (1983) 187, 205.
- [25] M. Lüscher and P. Weisz, Comm. Math. Phys. 97 (1985) 59; 98 (1985) 433 (E).
- [26] M. Lüscher and P. Weisz, Phys. Lett. B158 (1985) 250.
- [27] M. Lüscher and P. Weisz, Nucl. Phys. B266 (1986) 309.
- [28] M. Lüscher et al., hep-lat/9608049, to appear in the proceedings of Lattice'96, St. Louis, USA, 4-8 June 1996.

- [29] G.P. Lepage and P.B. Mackenzie, Phys. Rev. D48 (1993) 2250.
- [30] G. Parisi, in *High Energy Physics-1980*, eds. L. Durand and L.G. Pondrom (American Institute of Physics, New York, 1981).
- [31] M. Alford et al., Phys. Lett. B361 (1995) 87.
- [32] V. Periwal, Phys. Rev. D53 (1996) 2605.
- [33] J. Groeneveld et al., Physica Scripta 23 (1981) 1022; A. González-Arroyo and M. Okawa, Phys. Rev. D27 (1983) 2397.
- [34] P. van Baal, *Twisted boundary conditions: a non-perturbative probe for pure non-abelian gauge theories* (Ph.D. thesis, Utrecht, July 1984).
- [35] A. González-Arroyo and C.P. Korthals Altes, Nucl. Phys. B311 (1988) 433.
- [36] C.G. Callen et al., Phys. Rev. D17 (1978) 2717.
- [37] A.S. Kronfeld, Nucl. Phys. B (Proc. Suppl.) 4 (1988) 329; M. Teper, Nucl. Phys. B (Proc. Suppl.) 20 (1990) 159.
- [38] M. Lüscher, Comm. Math. Phys. 85 (1982) 39; Nucl. Phys. B200[FS4] (1982) 61.
- [39] M. Teper, Phys. Lett. B162 (1985) 357; Phys. Lett. B171 (1986) 81, 86.
- [40] Y. Iwasaki and T. Yoshié, Phys. Lett. B131 (1983) 159.
- [41] G. Savvidy, Phys. Lett. B71 (1977) 133; N. Nielsen and P. Olesen, Nucl. Phys. B144 (1978) 376; T. Hansson et al., Phys. Rev. D26 (1982) 2069; T. Hansson et al., Nucl. Phys. B289 (1987) 682.
- [42] B. Berg, Phys. Lett. B104 (1981) 475; J. Hoek et al., Nucl. Phys. B288 (1987) 589.
- [43] G. 't Hooft, Comm. Math. Phys. 81 (1981) 267; P. van Baal, Comm. Math. Phys. 94 (1984) 397.
- [44] P.J. Braam and P. van Baal, Comm. Math. Phys. 122 (1989) 267.
- [45] For a review see T. Eguchi et al., Phys. Rep. 66 (1980) 213.
- [46] M. García Pérez et al., Phys. Lett. B235 (1990) 117; M. García Pérez and A. González-Arroyo, J. Phys. A26 (1993) 2667; M. García Pérez et al., Phys. Lett. B305 (1993) 366.
- [47] S. Sedlacek, Comm. Math. Phys. 86 (1982) 515.
- [48] P. Braam et al., Inv. Math. 108 (1992) 419.
- [49] J. Ambjørn and H. Flyvbjerg, Phys. Lett. B97 (1980) 241.
- [50] P. van Baal, Comm. Math. Phys. 92 (1983) 1.
- [51] P. van Baal, Phys. Lett. B140 (1984) 375.
- [52] F. Bogomolov and P. Braam, Comm. Math. Phys. 143 (1992) 641.
- [53] C. Taubes, J. Diff. Geom. 19 (1984) 517.
- [54] S. Donaldson and P. Kronheimer, *The geometry of four manifolds* (Oxford University Press, 1990); D. Freed and K. Uhlenbeck, *Instantons and four-manifolds*, M.S.R.I. publ. Vol. I (Springer, New York, 1984).
- [55] P. van Baal, Theorem 4 in the set of theorems appended to the Ph.D. thesis *Twisted boundary conditions: a non-perturbative probe for pure non-abelian gauge theories* (Utrecht, July 1984).
- [56] J.A.M. Vermaseren, *Symbolic manipulation with FORM* (CAN, Amsterdam, 1992).
- [57] I. Are'feva, Teoreticheskaya i Matematicheskaya Fizika 43 (1980) 111 (translation p. 353).
- [58] N. Manton, Phys. Lett. B96 (1980) 328.
- [59] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* p.374 (Dover publ., New York, 1972).

- [60] L.S. Brown et al., Phys. Lett. B71 (1977) 103.
- [61] S. Wolfram et al., *Mathematica* (Addison-Wesley, New York, 1991).
- [62] M. García Pérez et al., Nucl. Phys. B (Proc. Suppl.) 34 (1994) 228; M. García Pérez and P. van Baal, Nucl. Phys. B429 (1994) 451.
- [63] Ph. de Forcrand et al., Nucl. Phys. B (Proc. Suppl.) 47 (1996) 777; hep-lat/9608032, to appear in the proceedings of Lattice'96, St. Louis, USA, 4-8 June 1996.
- [64] A.A. Belavin and A.M. Polyakov, J.ETP-Lett. 22 (1975) 245.
- [65] A.M. Polyakov, Phys. Lett. B59 (1975) 79; E. Brézin and J. Zinn-Justin, Phys. Rev. B14 (1976) 3110.
- [66] J.-L. Richard and A. Rouet, Nucl. Phys. B211 (1983) 447.
- [67] C. Michael and P.S. Spencer, Phys. Rev. D50 (1994) 7570.
- [68] E. Mottola and A. Wipf, Phys. Rev. D39 (1989) 588.
- [69] L. Alvarez-Gaumé et al., Ann. Phys. (N.Y.) 134 (1981) 85.
- [70] M. Green et al., *Superstring theory*, Vol.1 (Cambridge University Press, 1987).
- [71] A. D'Adda et al., Nucl. Phys. B146 (1978) 63.
- [72] V.A. Fateev et al., Nucl. Phys. B154 (1979) 1.
- [73] B. Berg and M. Lüscher, Comm. Math. Phys. 69 (1979) 57.
- [74] M. Blatter et al., Phys. Rev. D53 (1996) 923.
- [75] M. Lüscher and P. Weisz, Nucl. Phys. B240[FS12] (1984) 349.
- [76] M. García Pérez et al., hep-lat/9608015, to appear in the proceedings of Lattice'96, St. Louis, USA, 4-8 June 1996; M. García Pérez et al., Phys. Lett. B389 (1996) 112.
- [77] R. Dashen and D.J. Gross, Phys. Rev. D23 (1981) 2340.
- [78] P. Weisz and R. Wohlert, Nucl. Phys. B236 (1984) 397; B247 (1984) 544 (E).
- [79] P. Weisz, Nucl. Phys. B212 (1983) 1.
- [80] W. Celmaster and R. Gonsalves, Phys. Rev. D20 (1979) 1420.
- [81] A. Hasenfratz and P. Hasenfratz, Phys. Lett. B93 (1983) 165.
- [82] P. Weisz, Phys. Lett. B100 (1981) 331.
- [83] *The NAG Fortran Library Introductory Guide* (NAG, Oxford, 1993).
- [84] M. García Pérez and P. van Baal, hep-lat/9610036, Phys. Lett. B, in press.
- [85] C.T. Fike, *PL/I for scientific programmers* (Prentice Hall, 1970).
- [86] M. Lüscher and P. Weisz, private communication.
- [87] P. van Baal, private communication.



## List of publications

- M. García Pérez, A. González-Arroyo, J. Snippe and P. van Baal,  
*Instantons from over-improved cooling*,  
Nucl. Phys. B413 (1994) 535.
- M. García Pérez, A. González-Arroyo, J. Snippe and P. van Baal,  
*On the top of the energy barrier*,  
Nucl. Phys. B (Proc. Suppl.) 34 (1994) 222.
- J. Snippe,  
*Tunneling through sphalerons: the  $O(3)$   $\sigma$ -model on a cylinder*,  
Phys. Lett. B335 (1994) 395.
- J. Snippe and P. van Baal,  
*A new approach to instanton calculations in the  $O(3)$  nonlinear  $\sigma$ -model*,  
Nucl. Phys. B (Proc. Suppl.) 42 (1995) 779.
- M. García Pérez, J. Snippe and P. van Baal,  
*Testing improved actions*,  
to appear in the proceedings of the 2<sup>nd</sup> Workshop on Continuous Advances in  
QCD, Minneapolis, USA, 28–31 March 1996.
- M. García Pérez, J. Snippe and P. van Baal,  
*Improved action and Hamiltonian in finite volumes*,  
to appear in the proceedings of Lattice'96, St. Louis, USA, 4–8 June 1996.
- M. García Pérez, J. Snippe and P. van Baal,  
*A Monte Carlo study of old, new and tadpole improved actions*,  
Phys. Lett. B389 (1996) 112.
- J. Snippe,  
*Square Symanzik action to one-loop order*,  
Phys. Lett. B389 (1996) 119.
- J. Snippe,  
*Computation of the one-loop Symanzik coefficients for the square action*,  
in preparation.

## List of publications

1. ...  
2. ...  
3. ...  
4. ...  
5. ...  
6. ...  
7. ...  
8. ...  
9. ...  
10. ...  
11. ...  
12. ...  
13. ...  
14. ...  
15. ...  
16. ...  
17. ...  
18. ...  
19. ...  
20. ...  
21. ...  
22. ...  
23. ...  
24. ...  
25. ...  
26. ...  
27. ...  
28. ...  
29. ...  
30. ...  
31. ...  
32. ...  
33. ...  
34. ...  
35. ...  
36. ...  
37. ...  
38. ...  
39. ...  
40. ...  
41. ...  
42. ...  
43. ...  
44. ...  
45. ...  
46. ...  
47. ...  
48. ...  
49. ...  
50. ...  
51. ...  
52. ...  
53. ...  
54. ...  
55. ...  
56. ...  
57. ...  
58. ...  
59. ...  
60. ...  
61. ...  
62. ...  
63. ...  
64. ...  
65. ...  
66. ...  
67. ...  
68. ...  
69. ...  
70. ...  
71. ...  
72. ...  
73. ...  
74. ...  
75. ...  
76. ...  
77. ...  
78. ...  
79. ...  
80. ...  
81. ...  
82. ...  
83. ...  
84. ...  
85. ...  
86. ...  
87. ...  
88. ...  
89. ...  
90. ...  
91. ...  
92. ...  
93. ...  
94. ...  
95. ...  
96. ...  
97. ...  
98. ...  
99. ...  
100. ...

## Samenvatting

'Toepassingen van Verbeterde Acties in Roosterijktheorie' is de Nederlandse titel van dit proefschrift. In deze samenvatting zullen we uitleggen wat 'roosterijktheorie' is, en hoe die theorie past in het natuurkundige beeld op de wereld, in het bijzonder zoals beschreven door de hoge-energiefysica. In de roosterijktheorie is 'verbeterde actie' een uitermate natuurlijk begrip. Dat zullen we verderop ook uitleggen. Daarna geven we kort weer hoe in dit proefschrift verbeterde acties geconstrueerd, toegepast en getest worden.

Hoge-energiefysica is de tak van de natuurkunde die zich bezighoudt met de kleinste bekende deeltjes. Om deze deeltjes via een experiment goed te kunnen bestuderen, moet men ze met veel energie, dus met hoge snelheid, op elkaar laten botsen. Dat gebeurt in deeltjesversnellers, zoals op het CERN in Genève.

Het is de taak van de theoretische natuurkunde om experimentele waarnemingen te begrijpen (en te voorspellen) door middel van een model. Quantummechanica vormt de basis van kleine-deeltjesmodellen. Maar als de deeltjes bijna zo snel gaan als licht, zoals in versnellers, dan moet de speciale relativiteitstheorie van Einstein ingepast worden. Deze inpassing heeft geresulteerd in zogenaamde veldentheorieën. Een veld is iets dat in ieder punt van de ruimte een waarde heeft. Daardoor spreiden velden zich in principe over de hele ruimte uit, en is het in eerste instantie niet duidelijk hoe ze deeltjes kunnen beschrijven. Wel kunnen er in velden allerlei structuren optreden. De simpelste structuur is een golf. Bekende voorbeelden zijn golven in het elektromagnetische veld, die gebruikt worden voor de verzending van radio- en TV-signalen.

Als we quantummechanische effecten meenemen, blijkt dat golven uit kleine pakketjes zijn opgebouwd. Zulke pakketjes identificeren we met deeltjes. In het geval van het elektromagnetische veld zijn dit de zogenaamde fotonen. Om andere soorten deeltjes te kunnen beschrijven, moeten we andere velden introduceren. Overigens is het soms zinnig om de quantummechanische effecten buiten beschouwing te laten. Dit noemt men de klassieke limiet.

Het mooie is dat slechts één veld voldoende is voor de beschrijving van een willekeurig aantal deeltjes (van de soort behorende bij dat veld). Helaas moet hiervoor een prijs betaald worden: veldentheorieën zijn behoorlijk gecompliceerd. Dit komt doordat, per definitie, velden in ieder punt van de ruimte een andere waarde mogen hebben. Aangezien de ruimte uit oneindig veel punten bestaat, bezit een veldentheorie dus oneindig veel vrijheidsgraden. Dat maakt berekeningen meestal erg lastig. Bovendien blijkt er quantummechanisch een groot gevaar te zijn dat al die vrijheidsgraden een zodanig cumulatief effect teweeg brengen, dat de theorie geen zinnige voorspellingen oplevert. Slechts weinig veldentheorieën ontsnappen aan dit gevaar. Deze worden renormaliseerbaar genoemd.

Een voorbeeld van een renormaliseerbare veldentheorie is quantumelektrodynamica (QED), die elektronen en fotonen beschrijft. Deze theorie is opgebouwd uit een fotonveld (d.w.z. het elektromagnetische veld) en een elektronveld. Deze velden vertonen wisselwerking met elkaar. De elektromagnetische kracht (tussen elektronen) wordt zodoende begrepen als het uitwisselen (door elektronen) van fotonen.

We weten dus hoe elektronen en fotonen beschreven moeten worden. Er zijn echter nog meer deeltjes. De bekendste zijn protonen en neutronen, de bouwstenen van atoomkernen. Hiervoor ligt de zaak wat ingewikkelder. Terwijl er geen enkele aanwijzing is dat elektronen uit nog kleinere deeltjes zouden bestaan, staat het ondertussen wel vast dat protonen (en neutronen) uit drie quarks opgebouwd zijn. Voor een goed begrip van het proton is het daarom nodig om niet een protonveld in te voeren, maar quarkvelden. Verder is het zo dat de elektromagnetische kracht zowel qua sterkte als qua structuur niet de oorzaak kan zijn van het bijeenhouden van de quarks in een proton. Daardoor moet een nieuwe kracht ingevoerd worden. Deze kracht is derhalve niet het gevolg van de uitwisseling van fotonen, maar van de uitwisseling van een nieuw soort deeltjes. Deze deeltjes hebben de naam gluonen (lijmdeeltjes) gekregen. Het bijbehorende veld heet simpelweg het gluonveld. De resulterende theorie van quarks en gluonen noemt men quantumchromodynamica (QCD). Voor de volledigheid vermelden we dat modellen zoals QED en QCD tot de klasse van ijktheorieën behoren. Deze terminologie slaat op de mooie wiskundige structuur waarin ze gegoten kunnen worden, maar het voert te ver om dit hier preciezer uiteen te zetten.

Er is iets grappigs aan de hand met QCD. Op quantummechanische gronden moeten de drie quarks die een proton opbouwen allemaal van een ander type zijn. We moeten dus drie quarkvelden invoeren. Verder is de kracht tussen de quarks afhankelijk van het quarktype. Daardoor zijn er ook meerdere gluonvelden nodig. Het blijkt dat de enige manier om de theorie renormaliseerbaar (zinnig) te houden is om gluonvelden van een verschillend type met elkaar te laten wisselwerken. Dit is heel anders dan in QED: twee fotonen vliegen (althans in de klassieke limiet) dwars door elkaar heen, maar twee gluonen verstoren elkaars beweging.

Dit heeft zeer verstrekkende gevolgen, niet alleen voor het gedrag van gluonen, maar ook voor dat van quarks. Terwijl elektronen als afzonderlijke deeltjes kunnen rondvliegen, zitten quarks aan elkaar vastgeklonken. Dit is het zogenaamde opsluitingsprincipe. Probeert men twee quarks uit elkaar te trekken, dan vormt zich tussen de quarks een soort elastiekje, opgebouwd uit gluonen. Hierdoor zou het onindig veel energie kosten om een quark los te trekken, wat dus onmogelijk is.

Het is niet goed bekend hoe de structuur van het gluon-elastiekje met pen en papier uitgerekend moet worden. Dat is jammer, want deze structuur, en de manier waarop die tot stand komt, kan veel inzicht verschaffen over de opbouw van andere protonen en neutronen. Gelukkig biedt de computer uitkomst. Het is zelfs zo, dat pas sinds de komst van computersimulaties het bovenstaande opsluitingsscenario boven elke redelijke twijfel verheven is. Overigens is het voor de bestudering van het gluon-elastiekje verleidelijk om de theorie wat te versimpelen door de indirecte

invloed van de quarks helemaal buiten beschouwing te laten. Dat is wat we in dit proefschrift doen, en wat ook in veel simulaties gedaan wordt. Het versimpelde model heet pure-ijktheorie. Voor het onderstaande verhaal is het onderscheid tussen QCD en pure-ijktheorie niet zo belangrijk.

Een goede simulatie van QCD is niet eenvoudig, want de oneindig veel vrijheidsgraden die een veldentheorie bezit, passen simpelweg niet in het geheugen van een computer. De uitweg is een lichte verandering van de eigenlijke theorie. Hiertoe maken we het aantal punten in de ruimte eindig. Ten eerste passen we een discretisatie toe. Dat wil zeggen dat we de ruimte in vierkante blokjes ophakken, en ons beperken tot de waarden van de velden in het centrum van ieder blokje. De centropunten vormen een rooster. Vervolgens beperken we ons tot een eindig totaal volume, zodat we in iedere ruimterichting kunnen volstaan met een eindig aantal blokjes naast elkaar. Nu hebben we een eindig aantal vrijheidsgraden. Wat we tenslotte moeten doen, is de manier waarop de vrijheidsgraden met elkaar wisselwerken te formuleren op een manier die compatibel is met de roosterstructuur. Dit gebeurt door de definitie van een zogenaamde (rooster)actie. Dat is een functie van alle vrijheidsgraden, die op een slimme manier alle mogelijke wisselwerkingen samenvat. De actie die standaard gekozen wordt, heet de Wilson actie.

Op deze manier krijgen we een roosterijktheorie. Het is duidelijk dat deze afwijkingen zal vertonen van de eigenlijke theorie. Ten eerste is er het eindige volume. De effecten hiervan zijn goed begrepen, en zijn verwaarloosbaar klein zolang het volume maar groot genoeg is om ruimschoots één proton te kunnen omvatten. De andere afwijking wordt veroorzaakt doordat de afmeting van de blokjes (de roosterconstante) groter is dan nul. Deze afwijking wordt slechts langzaam kleiner bij verkleining van de blokjes. Daardoor stuiten we op het volgende probleem: voor nauwkeurige simulaties zouden we de roosterconstante erg klein moeten kiezen, maar het aantal blokjes is dan zo groot dat het de computer teveel tijd kost om alles goed door te rekenen.

De oplossing ligt in het gebruik van verbeterde acties. Dit zijn acties die bij een gegeven blokjesgrootte een beter resultaat geven dan de standaard Wilson actie. Met 'beter' bedoelen we dat de afwijking ten opzichte van de echte theorie, zoals geformuleerd in het continuüm (d.w.z. bij roosterconstante nul), kleiner is.

In dit proefschrift houden we ons hoofdzakelijk bezig met het onderzoeken van dergelijke aangepaste roosteracties. Dit is nodig, omdat niet iedere voorgestelde methode om de actie te verbeteren in de praktijk werkelijk tot betere resultaten leidt. Het is bijvoorbeeld mogelijk dat de verbetering pas inzet bij kleinere roosterconstanten dan meestal in simulaties worden gebruikt.

In hoofdstuk 2 beschouwen we niet zozeer verbeterde acties, maar alternatieve acties meer in het algemeen. We verrichten een systematische analyse van de klassieke structuur van roosteracties, en gebruiken die analyse om acties met een 'doorgeschooten verbetering' te construeren. De afwijking van zulke acties is in grootte ongeveer gelijk aan die van de Wilson actie, maar in teken tegengesteld. Dit kan toegepast worden in computeralgoritmes voor het vinden van zogenaamde instantonen, en wel de grootste die in een zeker volume passen.

Instantonen vormen weer een onderwerp apart. Aangezien ze een gedeelte van de stof van dit proefschrift vormen, zullen we er toch iets over zeggen. Instantonen zijn net als golven structuren in het gluonveld, maar ingewikkelder van aard. Hoewel het klassieke dingen zijn, hebben ze toch invloed op het quantummechanische gedrag van de theorie. Ze hebben indirect namelijk te maken met één van de meest prominente effecten van de quantummechanica: tunnelen. Dit duidt op de mogelijkheid om aan de andere kant van een barrière te komen zonder dat er genoeg energie is om er overheen te gaan.

Wanneer men probeert om de gluon-elastiekjes via een gewone berekening te begrijpen in plaats van via de computer, dan wordt men al vlug gedwongen om instantonen in ogenschouw te nemen. Dat levert complicaties op, ten eerste omdat de relevante instantonen niet altijd bekend zijn. De zoektocht naar deze instantonen is dan ook de eigenlijke motivatie voor hoofdstuk 2. Ten tweede zijn niet zozeer de instantonen zelf van belang, maar veeleer de vorm van de bovengenoemde barrières. Zodra men de instantonen bepaald heeft, is het daarom belangrijk om hun relatie tot die barrièrevorm te kennen. Ter illustratie behandelt hoofdstuk 3 deze relatie in een veldentheorie die enigzins lijkt op QCD, maar veel eenvoudiger van aard is.

Vanaf hoofdstuk 4 storten we ons volledig op verbeterde acties. Onze aandacht gaat uit naar verbeteringsmethoden die gebaseerd zijn op de door Symanzik bedachte (en door Lüscher en Weisz uitgewerkte) aanpak. Deze vallen uiteen in twee klassen: de Symanzik-verbetering zelf, en de zogenaamde 'tadpole'-verbetering. De eerste is rigoreus, maar houdt niet of slechts gedeeltelijk rekening met quantummechanische effecten. Dit betekent helaas dat de Symanzik-verbetering pas bij tamelijk kleine roosterconstanten goed begint te werken. De tweede berust op een uitbreiding van de Symanzik-methode, en probeert quantummechanische effecten nauwkeuriger mee te nemen. De argumentatie die ten grondslag ligt aan de tadpole-methode is echter niet zo goed onderbouwd. Daarom is een kritische analyse wenselijk, zelfs al bestaan er studies die laten zien dat de tadpole-methode een aanzienlijke verbetering kan geven. In hoofdstuk 4 tonen we echter aan dat de tadpole-methode in sommige gevallen helemaal niet effectief is, wat betekent dat in het algemeen voorzichtigheid met het gebruik ervan geboden is.

In hoofdstuk 5 construeren we een alternatief voor de Symanzik-verbeterde actie van Lüscher en Weisz. Het is namelijk zo dat er in principe een enorme keuzevrijheid is voor de fabricage van verbeterde acties, al kost die fabricage in de praktijk al vlug veel werk. Zowel de Lüscher-Weisz actie als onze nieuwe actie zijn zogenaamde één-lus verbeterde acties, wat aangeeft dat ze in een bepaalde welgedefinieerde zin quantummechanische effecten één stap beter meenemen dan de meest simpele (nul-lus, ofwel klassiek) verbeterde acties van het Symanzik-type. Zo'n alternatief kan in simulaties gebruikt worden om te onderzoeken in hoeverre de effectiviteit van de Symanzik-methode van de gebruikte actie afhangt. Bovendien geeft onze berekening de mogelijkheid om de tadpole-methode nog eens te testen. Op grond van een nul-lus Symanzik-verbeterde actie geeft de tadpole-methode namelijk een voorspelling voor de corresponderende één-lus Symanzik-verbeterde actie. Het is opmerkelijk dat deze

voorspelling tamelijk goed klopt, zowel voor de Lüscher-Weisz actie als voor de onze.

Al met al reikt dit proefschrift methoden aan die gebruikt kunnen worden om pure-ijktheorie beter te onderzoeken. Op hoofdstuk 2 geïnspireerde computeralgoritmes zijn door anderen al toegepast. Daarnaast zijn de door ons gevonden instantonen mogelijk een inspiratie voor toekomstige berekeningen. Verder verwachten we dat onze nieuwe Symanzik-actie (hoofdstuk 5) in simulaties van nut zal blijken bij gebruik naast de Lüscher-Weisz actie. Tot slot kan de nieuwe actie op zichzelf ook van pas komen. Zo kent de nul-lus versie toepassingen bij eindige-temperatuurberekeningen.

The text on this page is extremely faint and illegible. It appears to be a summary or abstract of a document, but the specific content cannot be discerned. The text is organized into several paragraphs, with some lines appearing to be bolded or indented, but the details are lost due to the low contrast and blurriness of the scan.

## Curriculum Vitae

Ik ben op 23 februari 1969 geboren te Leusden. Middelbaar onderwijs volgde ik aan 'De Drie Waarden' te Schoonhoven, en vervolgens aan 'Het Wagenings Lyceum' te Wageningen waar ik in 1987 het VWO-diploma behaalde. Ik vertrok naar de Universiteit Utrecht voor een studie Natuurkunde. In juni 1988 slaagde ik voor het propedeutisch examen, en in augustus 1992 voor het doctoraal examen. Mijn afstudeerwerk betrof 'baryoncreatie in deeltjesversnellers' en vond plaats onder leiding van prof.dr. G. 't Hooft.

In november 1992 begon ik mijn werkzaamheden als Assistent In Opleiding bij het Instituut-Lorentz voor theoretische natuurkunde te Leiden. Bij prof.dr. P.J. van Baal verrichtte ik een promotieonderzoek, met dit proefschrift als resultaat. Daarnaast verzorgde ik in 1993 en 1994 het werkcollege quantumveldentheorie.

Ik heb meerdere hoge-energiefysicascholen bezocht: in 1993 de workshop te Kloster Banz (Duitsland), in 1994 en 1995 de AIO-winterschool te Dalfsen, en in 1994 de zomerschool te Les Houches (Frankrijk). Verder heb ik deelgenomen aan de *Lattice* conferenties te Amsterdam (1992), Bielefeld (Duitsland, 1994) en St. Louis (Verenigde Staten, 1996). Bij de laatste twee gelegenheden heb ik delen van mijn promotieonderzoek door middel van een voordracht gepresenteerd.

## Stellingen

1. Discretisatie-effecten bieden de mogelijkheid tot de constructie van efficiënte numerieke algoritmes voor de bepaling van continuümplossingen met een schaalparameter.

*Hoofdstuk 2 van dit proefschrift.*

2. Zij  $U = P \exp(\int_0^a ds \bar{A}(s))$  een padgeordende exponent. Er geldt dat  $U = \exp(aA)$  met

$$A = \bar{A}^{(0)} + \frac{1}{12} a^2 \left\{ \frac{1}{2} \bar{A}^{(2)} + [\bar{A}^{(0)}, \bar{A}^{(1)}] \right\} + \mathcal{O}(a^3),$$

waarin  $\bar{A}^{(n)} \equiv \left. \frac{d^n}{ds^n} \bar{A}(s) \right|_{s=\frac{a}{2}}$ .

*Hoofdstuk 5 van dit proefschrift.*

3. Voor een gladde ( $C^\infty$ ) periodieke functie  $f$  wordt de integraal  $\int_{-T/2}^{T/2} dt f(t)$  zeer goed benaderd door de simpele discretisatie  $a \sum_{n=0}^{N-1} f(an)$  (met  $a \equiv T/N$ ,  $N \in \mathbb{N}$ ): voor  $a \rightarrow 0$  verdwijnt de fout sneller dan  $a^k$  ( $k \in \mathbb{N}$  willekeurig). Dit geldt ook op een oneindig interval, mits de functie glad en integreerbaar is en voldoende net gedrag vertoont voor  $x \rightarrow \pm\infty$ . Daardoor levert de keuze  $a = 1$  bij  $f(x) = \exp(-x^2/2)$  het volgende imposante resultaat:

$$\left| 1 - \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} e^{-n^2/2} \right| < 6 \cdot 10^{-9}.$$

4. De vier regels boven de laatste regel in formule (A.20) van ref. [1] dienen vervangen te worden door 0.

[1] P. Weisz en R. Wohlert, *Nucl. Phys. B236 (1984) 397*.

5. Zij  $U_{\mu\nu}^{m,n}(x) \in SU(N)$  de Wilson-loop behorende bij de rechthoekige contour  $x \rightarrow x + ma\hat{\mu} \rightarrow x + ma\hat{\mu} + na\hat{\nu} \rightarrow x + na\hat{\nu} \rightarrow x$  ( $n, m \in \mathbb{N}$ ;  $a$  de roosterconstante). Definieer de klasse van acties ( $k \in \mathbb{N}$ )

$$S_k = \sum_{x, \mu, \nu} \sum_{n=1}^k (-1)^{n+1} \frac{2}{n^4} \frac{1 \cdot (1-1/k) \cdots (1-(n-1)/k)}{(1+1/k) \cdot (1+2/k) \cdots (1+n/k)} \text{Tr} \left( 1 - U_{\mu\nu}^{n,n}(x) \right).$$

Op klassiek niveau vertoont  $S_k$  slechts  $\mathcal{O}(a^{2k})$  afwijkingen ten opzichte van de standaard pure-ijk continuümacie.

6. In ref. [2] worden hogere-orde afgeleidetermen aan het niet-lineaire sigma-model toegevoegd ter bestudering van cutoff-effecten. Er wordt beweerd dat de extra termen factoriseren in de Euler-Lagrange vergelijkingen. Deze bewering is onjuist. Derhalve zijn de Belavin-Polyakov instantonen geen exacte oplossingen in dit aangepaste model.

[2] V. Branchina en J. Polonyi, *Nucl. Phys. B433 (1995) 99*.

7. De actie

$$S = a \sum_{n \in \mathbb{Z}} \frac{1}{2} \left\{ (1 + 4\alpha) \left( \frac{\varphi_{n+1} - \varphi_n}{a} \right)^2 - \alpha \left( \frac{\varphi_{n+2} - \varphi_n}{a} \right)^2 + \omega^2 \varphi_n^2 \right\}$$

beschrijft een harmonische oscillator in (met stapgrootte  $a$ ) gediscrèteerde tijd. Het spectrum, gedefinieerd via de padintegraal en de transfer-matrixmethode van ref. [3], wordt gegeven door  $(n, m \in \{0, 1, 2, \dots\})$

$$E_{n,m} = \Omega_-(\frac{1}{2} + n) + \Omega_+(\frac{1}{2} + m), \quad \sinh(\frac{1}{2}a\Omega_{\pm}) \equiv \frac{1}{2} \sqrt{\frac{1}{2\alpha} (1 \pm \sqrt{1 - 4\alpha a^2 \omega^2})}.$$

Voor  $a \rightarrow 0$  geldt dat  $\Omega_+ = \mathcal{O}(1/a)$ , terwijl  $\Omega_- = \omega + \mathcal{O}(a^4)$  mits  $\alpha = \frac{1}{12}$ .

[3] M. Lüscher en P. Weisz, *Nucl. Phys. B240*[FS12] (1984) 349.

8. Toevoeging van een anharmonische term  $a \sum_n \lambda \varphi_n^4$  aan de actie in de vorige stelling verandert de  $a$ -afhankelijkheid drastisch: de leidende-orde afwijking in het laagliggende spectrum wordt lineair in  $a$ . Expliciet geldt het volgende:

$$E_{n,0} \rightarrow E_{n,0} + \frac{3}{2} \lambda \left\{ \frac{1}{\tilde{\omega}_-^2} (\frac{1}{2} + n + n^2) - \frac{2a}{\omega_+ \tilde{\omega}_-} (\frac{1}{2} + n) + \frac{a^2}{2\omega_+^2} \right\} + \mathcal{O}(\lambda^2),$$

$$\omega_{\pm} \equiv \sqrt{1 - 4\alpha a^2 \omega^2} \sinh(a\Omega_{\pm}), \quad \tilde{\omega}_- \equiv \frac{\omega_-}{a}.$$

9. Artikel 19b van het promotiereglement, luidende "De promotor deelt de promovendus mede of hij de stellingen verdedigbaar acht.", neemt een consciëntieuze promotor onvoldoende in bescherming tegen een zichzelf ontkennende stelling.
10. Oefening in spreekvaardigheid zou het aanleren van de klassieke talen bespoedigen.
11. Uitspraken die beginnen met "Zonder mezelf op de borst te willen slaan" hebben meestal een zeer merkwaardig vervolg.

Jeroen Snippe  
19 februari 1997