BEYOND THE GRIBOV HORIZON IN THE FEMTO UNIVERSE

Pierre van Baal *

Institute for Theoretical Physics
Princetonplein 5, P.O.Box 80.006
3508 TA Utrecht, The Netherlands

ABSTRACT

We will outline the semi-analytic calculations of the low-lying energy spectrum for pure gauge theory in an intermediate cubic volume of up to about 1 Fermi in each direction. The viability of the method is strongly supported by the close agreement found with the results obtained from lattice Monte Carlo calculations, especially after we incorporate in our approach the effects of the lattice artifacts. This paper will analyse the situation for *arbitrary* gauge groups. It involves dividing configuration space into a number of coordinate patches together with the prescription of glueing these patches together. Thus some topologically non-trivial behaviour is incorporated, in a way which avoids the Gribov ambiguity.

INTRODUCTION

The last few years a method has been developed to calculate the low-lying continuum spectrum for pure gauge theories in intermediate volumes. This was based on the analytic calculation of the effective Hamiltonian for the zero momentum modes, as first derived by Lüscher [1], removing some technical obstacles observed by 't Hooft [2] in formulating non-abelian gauge theories on the torus. The torus geometry (more precisely its nontrivial first homology) implies an extra symmetry for pure gauge theories, related to the center of the gauge group, whose representations are labelled by the electric flux $\mathbf{e}$, where e_i is labelled by the integers modulo N for the gauge group SU(N). Thus the energy of one unit of 't Hooft type electric flux is the difference in energy between the ground states in the sectors with $\mathbf{e} = (1, 0, 0)$ and $\mathbf{e} = \mathbf{0}$. Choosing furthermore a symmetric torus (i.e. formulating the gauge theory on a cube with periodic boundary conditions) one has in addition the cubic symmetry group as a discrete remnant of the rotational symmetry, the latter being the relevant symmetry group if the volume of the torus is taken to infinity. The mass of a glueball in intermediate volumes is calculated as the difference in energy between an excited state and the ground state of the effective Hamiltonian, where the glueball is labelled by the rep-

* KNAW fellow

resentation to which the excited state belongs. The ground state is as usual in the singlet scalar representation A_1^+. The remaining representations are the doublet E, the triplets T_1, T_2 and the singlet A_2, each occurring with both parities. For SU(3), or any gauge group with complex representations, there is also the charge conjugation symmetry. Recently Vohwinkel [3] has been able to generalize the intermediate volume calculations to SU(3).

The finite volume serves both to introduce an infrared cutoff, nevertheless retaining the zero-momentum modes and as a control parameter for changing the strength of the interaction. Such a control parameter is lacking in an infinite volume, due to the classical scale invariance of the theory. In a finite volume, the renormalization of the coupling constant relates the size of the volume to the strength of the interactions through the use of the renormalization group, and asymptotic freedom allows us to do perturbation theory in sufficiently small volumes. In this perturbative expansion, obviously zero-momentum modes will dominate, but it is crucial that they remain interacting to lowest non-trivial order [1]. This is an essential feature of non-abelian theories and might be a crucial aspect of the infrared problems, which thus survives partly in the finite volume analysis.

The aim of the finite volume expansion is to try to incorporate increasingly more complicated dynamical effects in the calculation, which become important for larger volumes (i.e. stronger interactions). What essentially happens is that the wave functional starts to spread out over configuration space and sees more of the topology of this configuration space and of the potential energy. There is certainly a strong connection between these two aspects: tunnelling is associated with potential barriers separating different vacua, which are related by topologically nontrivial gauge transformations [4,5]. Choosing local (gauge fixed) coordinates in a neighbourhood of one of these vacua will run into coordinate (gauge) singularities at the other vacua, which can be associated to Gribov horizons [6]. This was first observed in the analogous situation of instantons on S^4 [7]. It is these aspects that this contribution will emphasize. As followed from the analysis of Singer [8], Gribov horizons occur due to the topological nontrivial nature of this configuration space. It was Nahm [9] who first suggested setting up field theory using different coordinate patches in configuration space, emphasizing that Gribov horizons are just coordinate singularities. We will show how to avoid these coordinate singularities by using different coordinate patches with transition functions that can essentially be described as gauge transformations. Incorporating one of these effects, which can be largely (but not completely) formulated within the zero-momentum sector allows one to understand the origin of the energy of electric flux and the low-lying glueball masses in volumes of up to 0.7 fermi (5 times the correlation length of the scalar glueball).

One problem many of the approximations involved in trying to understand the strong interactions have, is that QCD is usually too formidable to attack in full and comparison with experiment is not yet a realistic goal. Yet, without such a comparison we are basically lost, since we can not test our approximations. It was therefore essential in the development we are going to illustrate, that a completely independent check on our approximations was possible in terms of a comparison with the results obtained by lattice Monte Carlo simulations. So instead of comparing with full QCD experimental results we are, for the time being, helped out by a comparison with the results obtained by computer experiments for SU(2) and SU(3) gauge theories in intermediate volumes [10,11]. The reason this comparison is possible, is due to the fact that computer simulations are done on finite rectangular lattices with periodic boundary conditions at a finite value of the bare lattice coupling constant. This means that the physical volume is finite. Of course, since the continuum limit is not actually taken (due to computer limitations) one is still left with results that do depend on the number of lattice sites, called the lattice artifacts, but it turns out that these are surprisingly small, at least for SU(2) in intermediate volumes. Furthermore, we could actually correct for these lattice artifacts [12] and the excellent agreement lends the ultimate justification of our approach. The essential lesson we learn is that in order to understand these strong interactions we have to go beyond the Gribov horizon, by which we mean to say that wave functionals will spread out over such a large portion of configuration space, that the problem of Gribov horizons and copies can no longer be ignored.

Finally, let us mention that massless fermions have been included in the finite volume analysis, for both SU(2) and SU(3). Their influence can be quite intricate [13], as is evident for SU(3) with periodic (as opposed to anti-periodic) boundary conditions for the fermions. In that case charge conjugation (C) and parity (P) symmetry are spontaneously broken in small volumes, but CP is conserved. For SU(2) the intermediate volume analysis is also possible, especially by using the Monte Carlo technique on the effective Lagrangian for the zero-momentum gauge fields [14].

In the remainder, we first explain the perturbative expansion for gauge theories on a torus. We will describe this for an arbitrary simple, simply connected gauge group G, in view of the recent generalization to $SU(3)$ [3]. Then we indicate where the Gribov horizons show up in this analysis and why it is essential to be able to disentangle its relevance for the dynamics of the theory. We conclude with a discussion of the results, especially in the light of the comparison with Monte Carlo calculations.

THE CLASSICAL VACUUM

When solving a quantum problem in perturbation theory (to start with), whether in ordinary quantum mechanics or field theory, one first has to find the classical vacuum. For gauge theories on the torus (or for that matter any three dimensional manifold M with non-trivial first homology) this first step already requires care, because it will turn out that this classical vacuum contains infinitely finite dimensional connected components. Each such a component will turn out to be an orbifold. An orbifold is a compact space modulo the action of a finite group. The fixed points of this group action are called orbifold singularities.

Before we describe this more explicitly, closely following the discussion in ref. [15] (app. D), we can already outline the essential features of our analysis with the present information. Two different connected components are related by a homotopically non-trivial gauge transformation $g : M \to G$. These are gauge transformations with a non-trivial winding number P

$$P = \frac{1}{24\pi^2} \int_{T^3} \mathrm{Tr}((g(\mathbf{x}) dg(\mathbf{x})^{-1})^3) \tag{1}$$

and are related to instantons of topological charge P, which describe the tunnelling between the two classical vacua, through a potential barrier, which necessarily separates them. Therefore, as long as the coupling constant is not too large, the wave functional will be sufficiently localized in the neighbourhood of the classical vacua, to be able to neglect the overlap of the wave functional of one classical vacuum, with the wave functional belonging to a neighbouring classical vacuum. Since Gauss's law tells us that wave functionals are invariant under homotopically trivial gauge transformations, the wave functional can only depend on the homotopy type P of the gauge transformation. These form the abelian homotopy group $\pi_3(G)$, and wave functionals can be classified as irreducible representations of this group, which are parametrized by the parameter θ. This is in complete analogy with the Bloch momenta for a periodic potential problem. Hence for sufficiently small coupling constants one can neglected the θ dependence of the energies and the analysis can be restricted to wave functionals supported in the neighbourhood of one connected component of the vacuum, which we call a vacuum valley.

As mentioned before this vacuum valley has orbifold singularities. The singularities arise due to fixed points of a group action, and this finite group must be a remnant of the group of gauge transformations (it will turn out to be the Weyl group $\mathcal{W}_G$). Therefore, the fixed points are related to points where the stablizer of the gauge field in question (i.e. the subgroup that leaves it invariant) will change, as compared to neighbouring points. This implies that the potential in directions transverse to the vacuum valley will be wider at these orbifold singularities and thus these are the places where the wave functional tends to concentrate (they are the caustics in an analogue with optics). At extreme weak coupling the theory is described by expanding around the orbifold singularity where the potential is widest. However, we will find out that the orbifold singularities themselves are connected by symmetries. These symmetries are again part of the group of gauge transformations, but since they relate two different orbifold singularities, the relevant gauge transformations are homotopically non-trivial. They are the ones related to the first homology of the manifold M, with values in the center Z_G of the gauge group G, $H_1(M, Z_G)$ (we will assume M to be such that the torsion of the first homology group vanishes). The representations of this abelian homology group are labelled by 't Hooft's definition of electric flux. For each generating circle of M (each generator of $H_1(M, Z_G)$) there is one component of electric flux.

It is now time to be more specific. From now on we will assume $M = T^3$, which for convenience is considered as $T^3 = R^3/\Sigma$, where Σ is a three dimensional lattice with lattice vectors $L\mathbf{f}^I$. The overall scale is set by L (we can for example normalize the unit lattice cell to have volume L^3, or $\det(f) = (\mathbf{f}^1 \times \mathbf{f}^2) \cdot \mathbf{f}^3 = 1$). The gauge fields will satisfy periodic boundary conditions

$$A_k(\mathbf{x} + L\mathbf{f}^I) = A_k(\mathbf{x}) \tag{2}$$

and take values in the Lie algebra L_G of the group G. We wish to find the set of gauge potentials modulo homotopically trivial gauge transformations

$$[g]A_k(\mathbf{x}) = g(\mathbf{x})A_k(\mathbf{x})g(\mathbf{x})^{-1} - ig(\mathbf{x})\partial_k g(\mathbf{x})^{-1} \quad , \tag{3}$$

which minimize the classical potential

$$V(A) = \frac{1}{2g^2}\int_{T^3} d_3x \mathrm{Tr}(F_{ij}^2(\mathbf{x})) \quad , \tag{4}$$

$$F_{ij} = \partial_i A_j - \partial_j A_i + i[A_i, A_j] \quad . \tag{5}$$

The minimum is clearly $V(A) = 0$ and is obtained by the so-called flat connections, gauge fields for which the curvature vanishes ($F_{ij}(\mathbf{x}) = 0$).

The homotopy of the gauge functions is specified completely by P in eq. (1), if we insist $g(\mathbf{x})$ to be periodic. However, if all the fields transform trivially under the center Z_G of the gauge group the theory is also invariant under gauge transformations that are periodic up to an element of the center of the gauge group

$$g(\mathbf{x}+L\mathbf{f}^J)=\exp(-2\pi i\Theta_J)g(\mathbf{x}) \quad . \tag{6}$$

Here $i\Theta_J \in L_G$ is such that $\exp(2\pi i\Theta_J)\in Z_G$. These gauge transformations are the so-called twisted gauge transformations introduced by 't Hooft [2]. One can indeed verify that the periodicity of the vector potential is conserved under these gauge transformations. It is easy to write down explicit representatives of eq.(6):

$$g_{\boldsymbol{\Theta}}(\mathbf{x})=\exp(-2\pi i\mathbf{x}\cdot\boldsymbol{\Theta}/L) \quad , \tag{7}$$

where

$$\boldsymbol{\Theta}=\sum_J \Theta_J\mathbf{h}^J \quad . \tag{8}$$

Thus, Θ_J are the components of Θ with respect to the basis $\mathbf{h}^J$ of the dual lattice $\tilde{\Sigma}$, i.e. $\mathbf{f}^I\cdot\mathbf{h}^J=\delta_{IJ}$. Note that the gauge function $g_{\boldsymbol{\Theta}}$ is essentially abelian and that $P(g_{\boldsymbol{\Theta}})=0$. One can prove [16] that any gauge function g with the homotopy type specified by $P=P(g)\in\mathbb{Z}$ and by $\exp(-2\pi i\Theta_I)\in Z_G$ can be written as $g=g_1^P g_{\boldsymbol{\Theta}} g_0$, where g_0 is homotopically trivial and g_1 has winding number $P=1$ (using the property $P(g\hat{g})=P(g)+P(\hat{g})$).

We will now specify the flat connections. It is clear that, if we take the gauge fields to be independent of the coordinates and to take values in the abelian Cartan subalgebra H_G, the curvature will vanish. It is not difficult to prove that any flat connection is of this form up to a periodic gauge transformation (e.g. see ref. [15], app. D). Thus we can parametrize the vacuum valley by the zero-momentum abelian connections

$$A_k(\mathbf{x})=\frac{1}{L}c_k^aT_a=\frac{1}{L}h_k^Ic_I^aT_a \quad . \tag{9}$$

We parametrized the Cartan subalgebra by the abelian generators T_a, where the index a ranges from 1 to the rank ℓ of the group G. Also note that we distinguish between the vector indices $i,j,\cdots$ and $I,J,\cdots$, where the latter indicate the components of a vector with respect to the basis of $\tilde{\Sigma}$. There is a canonical way of completing the generators T_a to a (complex) basis for the Lie algebra L_G, by diagonalizing the action of $\mathrm{ad}(H_G)$ (with $\mathrm{ad}(X)Y=[X,Y]$ for $X,Y\in L_G$, see [17] for mathematical background).

$$[T_a,E_\alpha]=\alpha_aE_\alpha \quad . \tag{10}$$

The real eigenvalues α_a form the root vectors α and they span a ℓ dimensional root lattice Λ_r. The inner product on the Euclidean space spanned by the roots is given by

$$(\alpha,\beta)=\alpha_ag^{ab}\beta_b \quad , \quad (g^{ab})^{-1}=g_{ab}=\mathrm{Tr}_{\mathrm{ad}}(T_aT_b) \quad . \tag{11}$$

Note that with α a root also $-\alpha$ is a root and that $E_{-\alpha}=E_\alpha^\dagger$.

The condition that $\exp(-2\pi i\Theta_J)\in Z_G$ clearly implies that $\Theta_J=\Theta_J^aT_a$ and $\exp(2\pi i\Theta_J^a\alpha_a)=1$ (from $\exp(-2\pi i\Theta_J)E_\alpha\exp(2\pi i\Theta_J)=E_\alpha$ for any root α). Thus Θ_J is an element of the lattice dual to the root lattice (this is called the dual weight lattice $\tilde{\Lambda}_w$). It can be explicitly defined by picking among the roots a basis $\Delta=$

$\{\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(\ell)}\}$, the dual weight lattice is then spanned by the vectors $\tilde{\lambda}_{(i)}$, which satisfy $\tilde{\lambda}^a_{(i)}\alpha^{(j)}_a = \delta_{ij}$.

It is now crucial to observe that under the gauge transformation with $g_{\boldsymbol{\Theta}}$, c^a_I will transform by a translation

$$[g_{\boldsymbol{\Theta}}]c^a_I = c^a_I + 2\pi\Theta^a_I \quad . \tag{12}$$

Expressing c^a_I with respect to the basis $\tilde{\lambda}_{(i)}$,

$$c^a_I = c^{(i)}_I \tilde{\lambda}^a_{(i)} \quad , \tag{13}$$

reduces the action of $[g_{\boldsymbol{\Theta}}]$ to a simple translation over 2π

$$[g_{\boldsymbol{\Theta}}]c^{(i)}_I = \alpha^{(i)}_a(c^a_I + 2\pi\Theta^a_I) = c^{(i)}_I + 2\pi n^{(i)}_I \quad . \tag{14}$$

Note that the $n^{(i)}_I$ are simply the integer components of Θ_I with respect to the basis for the dual weight lattice.

To obtain the set of inequivalent configurations, we still have to divide out the homotopically trivial gauge transformations that leave the set of zero momentum abelian gauge fields invariant. There are two types of transformations we have to consider. First $g_{\boldsymbol{\Theta}}$ with $\exp(-2\pi i\Theta_J) = 1$ and second the constant gauge transformations that leave the Cartan subalgebra invariant. This subgroup is called the normalizer of the Cartan subgroup H_G

$$N_G(H_G) = \{g \in G | ghg^{-1} \in H_G, \forall h \in H_G\} \quad . \tag{15}$$

The first set of transformations implies that there is a sublattice $\tilde{\Lambda}_r$ of the dual weight lattice such that for $\tilde{\Theta}_J \in \tilde{\Lambda}_r$, $\exp(-2\pi i\tilde{\Theta}_J) = 1$. One can show that $\tilde{\Lambda}_r$ is the dual root lattice, i.e. the lattice spanned by the dual roots $\tilde{\alpha} = 2\alpha/(\alpha,\alpha)$, using the fact that $\tilde{\alpha}^a T_a/2$ is the generator of an SU(2) subgroup (see eq. (10)). Therefore, we have the following equivalence relation,

$$c_I \sim c_I + 2\pi k_{I(i)}\tilde{\alpha}^{(i)} \quad , \quad k_{I(i)} \in \mathbb{Z} \quad . \tag{16}$$

The second set of transformations will generate automorphisms on the roots. Since the centralizer

$$C_G(H_G) = \{g \in G | ghg^{-1} = h, \forall h \in H_G\} \quad , \tag{17}$$

is a subgroup of the normalizer $N_G(H_G)$, which leaves the generators T_a fixed, the group of automorphisms of the roots is isomorphic to $N_G(H_G)/C_G(H_G)$. This group of automorphisms can be shown to be the Weyl group $\mathcal{W}_G$, which is generated by the Weyl reflections [17]

$$\sigma_\alpha(\beta) = \beta - 2(\beta,\alpha)\alpha/(\alpha,\alpha) \quad . \tag{18}$$

The Weyl group is a finite group and for SU(N) it is isomorphic to the group of permutations of N objects, $\mathcal{S}_N$. Hence, the vacuum valley is given by the orbifold

$$\mathcal{V} = [R^\ell/2\pi\tilde{\Lambda}_r]^3/\mathcal{W}_G \quad . \tag{19}$$

The symmetry related to $H_1(T^3, Z_G)$ is now realized on the vacuum valley $\mathcal{V}$ by the translations of c_I over $\Theta_I \in \tilde{\Lambda}_w/\tilde{\Lambda}_r$ and this group is isomorphic to Z^3_G. This symmetry acts non trivially on the wave functionals, since its origin is in the homotopically non trivial gauge transformations. The wave functionals have to be representations of this symmetry

$$[g_{\mathbf{\Theta}}]\Psi(A)=\exp(2\pi i e_a^J\Theta_J^a)\Psi(A) \quad , \tag{20}$$

where $\mathbf{e}$ has to satisfy the property that $e_a^J\tilde{\Theta}_J^a\in\mathbb{Z}$. This means that e_a^J is on the lattice dual to the dual root lattice $\tilde{\Lambda}_r$, which is precisely the weight lattice Λ_w. Furthermore, $e_a^J\Theta_J^a\in\mathbb{Z}$ for any $\Theta_J\in\tilde{\Lambda}_w$ implies that $e^J\in\Lambda_r$, where the root lattice Λ_r is indeed a sublattice of Λ_w. Thus the 't Hooft electric flux is an element of $(\Lambda_w/\Lambda_r)^3$, which is isomorphic to $Z^3_{\tilde{G}}$, where $\tilde{G}$ is the simply connected group dual to G (its root lattice is dual to the root lattice of G). For SU(N) the root lattice is self dual, such that $\tilde{G}=\mathrm{SU(N)}$. We end our extensive discussion of the vacuum valley by studying the fixed points of the Weyl group, which will form the orbifold singularities. Because of the symmetries induced by $g_{\mathbf{\Theta}}$ the fixed point set is the orbit under this shift symmetry of the fixed point set for the space spanned by $\tilde{\Lambda}_w$, which is easily seen to only contain the origin. The orbifold singularities therefore form a Z^3_G orbit of the origin. For SU(N) there are N^3 orbifold singularities. As we shall see in more detail these are the degenerate quantum vacua in perturbation theory.

THE PERTURBATIVE EXPANSION

As was already discussed in the context of the orbifold singularities, the width of the potential in directions transverse to the vacuum valley depends on the position in this valley. The most obvious way to proceed is thus to use the Born-Oppenheimer approximation of integrating out the fast fluctuations (fluctuations for which the potential energy changes), to be left with an effective theory for the slow fluctuations (those for which the potential doesn't change, which are hence tangent to the vacuum valley). To lowest order this amounts to a one-loop approximation of the effective Lagrangian. However, it is here that the orbifold singularities will demonstrate their singular nature. At these points, due to the invariance of this configuration under a bigger subgroup of the gauge group, there will be a few directions in configuration space, in which the potential varies quartic rather than quadratic in the fields. In a one-loop expansion around the orbifold singularity, these should therefore also be considered as slow modes, despite the fact that they are not tangent to the vacuum valley. This perturbative expansion was performed by Lüscher [1], by deriving an effective Hamiltonian for the zero-momentum modes. He also realized that to all orders in perturbation theory the energy of electric flux will be zero, because of an induced potential barrier separating the orbifold singularities.

Let us first assume no problems near the singularities of the orbifold, then to one-loop order the effective potential is the sum of the zero-point energies for the fluctuations transverse to the vacuum valley:

$$V_{eff}(c)=\frac{2}{L}\sum_\alpha\sum_{\mathbf{k}}||2\pi\mathbf{k}+\mathbf{c}_\alpha|| \quad , \tag{21}$$

where

$$\mathbf{k}=k_I\mathbf{h}^I \quad , \quad k_I\in\mathbb{Z} \quad , \quad \mathbf{c}_\alpha=\sum_{a=1}^{\ell}\mathbf{c}^a\alpha_a \quad . \tag{22}$$

The sum over α is over all roots. We refer to [15] for details. By using the heat-kernel and Poisson resummation [1,15] we can rewrite this up to a field independent term as

$$V_{eff}(c)=\frac{2\det(f)}{\pi^2L}\sum_\alpha\sum_{\mathbf{n}\neq\mathbf{0}}\frac{\sin^2(\mathbf{n}\cdot\mathbf{c}_\alpha/2)}{(\mathbf{n}\cdot\mathbf{n})^2} \quad , \quad \mathbf{n}=n_I\mathbf{f}^I \quad , \quad n_I\in\mathbb{Z} \quad . \tag{23}$$

As it should be, V_{eff} is invariant with respect to the Z_G^3 symmetry induced by the gauge transformations $g_{\mathbf{\Theta}}$. Furthermore, the latter form allows us to derive a useful approximation for V_{eff}, when restricted to one dimension, say $c_I^a = \delta_{1I} c^a$:

$$V_{eff}(c) = \frac{||\mathbf{h}^1||}{2\pi L}\{(4\pi\delta)^2 - (c - 4\pi\delta)^2 + 2\sum_{n=1}^{\infty}\sum_{\alpha} a_n(f)\sin^2(nc_\alpha/2)\} \quad , \quad c_\alpha \in [0, 2\pi] \quad , \tag{24}$$

where δ is the highest weight (also equal to half the sum of the positive roots)

$$\delta = \sum_{i=1}^{\ell} \lambda^{(i)} \quad , \quad (\lambda^{(i)}, \tilde{\alpha}^{(j)}) = \delta_{ij} \quad , \tag{25}$$

i.e., $\lambda^{(i)}$ form a basis for the weight lattice Λ_w. The squared vectors occurring in eq. (24) are defined through eq. (11). The constants $a_n(f)$, computed in [15] for a cubic lattice Σ ($f_k^I = \delta_k^I$) are small, except for very asymmetric lattices Σ (see [12]). Neglecting them in first approximation gives for SU(3) a simple picture of the potential as depicted in figure 1 by equipotential lines.

Figure 1: This is a one dimensional cross section of the vacuum valley for SU(3), with $c = c^1\lambda_8/2 + c^2\lambda_3/2 = x\lambda_8/2 + y\lambda_3/2$. The quantum vacua are indicated by 0 and $\pm$, corresponding to $\exp(-2\pi i\Theta) = 1$ respectively $\exp(\pm 2\pi i/3)$. The circles are equipotential lines for $V_{eff}(c)$ at equidistant values of the potential, with its maximum at $4\pi\delta$. The shaded area is a Weyl chamber, parametrizing the gauge invariant ("radial") coordinates (r,s) and the fat hexagon is the Gribov horizon for the perturbative expansion around the origin. The Hamiltonian theory can be entirely defined on the coordinate patch indicated by the dashed hexagon, with appropriate boundary conditions.

This figure is generic for arbitrary gauge group and it allows us to analyse the nature of the orbifold singularities in a more instructive way. These singularities are clearly visible as the conic or V-shape singularities in the equipotential lines. They are easily seen to arise due to the vanishing of one or more of the quadratic fluctuations. For example, when $\mathbf{k}=\mathbf{0}$, $\mathbf{c}_\alpha=\mathbf{0}$ for any combination of the α explains the singularities emanating from the origin. They are the result of the invariance of these configurations under a subgroup of the gauge group G. This can be seen from the fact that the planes fixed under a Weyl reflection are precisely spanned by the vectors of the dual weight lattice, e.g. $\sigma_{\alpha^{(i)}}(\tilde{\lambda}_{(j)})=\tilde{\lambda}_{(j)}$ for $i\neq j$. These planes are also known as the boundaries of the Weyl chambers [17]. The point singularities are invariant under the whole group, verifying that the quantum vacua ($\mathbf{c}_\alpha=\mathbf{0}$ for all α and their Z_G^3 conjugates) are fixed points of the Weyl group. Note that for SU(2) only point singularities arise, this is one of the reasons why SU(2) is not entirely generic.

We now have to re-examine the Born-Oppenheimer approximation, due to the vanishing of some of the eigenvalues of the transverse fluctuations, which will invalidate the separation of the slow and fast modes (it gives rise to the breakdown of the adiabatic approximation as discussed in [4,15]). As mentioned before, one has to include the modes associated to the vanishing eigenvalues for the transverse fluctuations in the slow modes. Expanding around the origin the slow modes are thus the modes with zero momentum for which Lüscher [1] derived an effective Hamiltonian. This is done in a perturbative expansion, starting from the Hamiltonian formulation [18] of gauge theory ($A_o=0$) in the coulomb gauge $\partial_i A_i=0$, using Bloch's method [19] to deal with degenerate perturbation theory. Alternatively, this result can be reproduced by deriving an effective Lagrangian [4,15], using the gauge $\partial_\mu A^\mu+i[\mathcal{P}A_\mu,A^\mu]=0$ (where $\mathcal{P}A_\mu=\int_{T^3}d_3x A_\mu(x)/(\det(f)L^3)$ and $\mathcal{P}A_o=0$) and integrating out the non-zero momentum modes.

The result can be written in lowest order as

$$
\begin{aligned}
H_{eff}(c) &= -\frac{g^2(L)}{2L\det(f)}\frac{\partial^2}{\partial c_i^a\partial c_i^a}+\frac{\det(f)}{4Lg^2(L)}F_{ij}^aF_{ij}^a+V_1(c) \quad , \\
V_1(c) &= \frac{1}{L}\sum_{\mathbf{k}\neq 0}\mathrm{Tr}_{\mathrm{ad}}\sqrt{(2\pi\mathbf{k}\mathbb{I}+\mathbf{c}^a\mathbb{T}_a)^2} \quad , \qquad (26)
\end{aligned}
$$

where $\mathbb{I}$ is the unit matrix, $\mathbb{T}_a$ are the generators in the adjoint representation, and $c_i^a=2\mathrm{Tr}(T_a\mathcal{P}A_i))$. Also, the index a runs over all values $1,2,\cdots,\mathrm{dim}G$ and we assumed the generators to be normalized according to $\mathrm{Tr}(T_aT_b)=\delta_{ab}/2$. Note that the sum over $\mathbf{k}$ in the definition of $V_1(c)$ excludes $\mathbf{k}=\mathbf{0}$, which is the mode responsible for the conic singularities discussed earlier. We have written it such that on the vacuum valley $V_{eff}(c)=2\sum_\alpha||\mathbf{c}_\alpha||/L+V_1(c)$. Along the vaccum valley the potential part of H_{eff} is therefore exact to one-loop order. Finally, $g(L)$ is the renormalized coupling constant at the scale L for the continuum, most conveniently calculated in the MS-scheme of dimensional regularization [1].

THE GRIBOV HORIZON AND BEYOND

If we now wish to extend this effective Hamiltonian to the whole vacuum valley we clearly will run into problems, since we only eliminated the singularities emanating from the origin, corresponding to the boundaries of the Weyl chambers. In particular, the Z_G^3 conjugates of the origin will still cause conic singularities in $V_1(c)$, which can be understood as arising due to the invariance of these points under the subgroup $g_{\mathbf{\Theta}} G g_{\mathbf{\Theta}}^{-1}$, which are in general not constant gauge transformations. It allows us to almost trivially associate them with Gribov horizons [6], which is the place where Gribov copies first form and where the gauge orbit has a direction tangent to the coordinate slice defined by the gauge condition. To be specific, let us therefore consider the Coulomb gauge $\partial_i A_i = 0$. We note that all configurations in the vacuum valley, $\mathbf{A} = \mathbf{c}/L$, satisfy the Coulomb gauge condition. The Faddeev-Popov determinant is given by

$$\rho(A) = \det{}'(FP(A)) \quad , \quad FP(A) = -\partial_i D_i(A) = -[\partial_i^2 + i\mathrm{ad}(A_i)\partial_i] \quad , \tag{27}$$

where the prime denotes that the determinant does not include the zero-modes arising as a consequence of the invariance of the gauge condition under constant gauge transformations.

When A is in the vacuum valley, $A_i = \sum_{a=1}^{\ell} c_i^a T_a / L$, we can actually diagonalize the operator $FP(A)$ in eq. (27)

$$\begin{aligned} FP(A) \exp(2\pi i \mathbf{k} \cdot \mathbf{x}/L) T_a &= (\frac{2\pi \mathbf{k}}{L})^2 \exp(2\pi i \mathbf{k} \cdot \mathbf{x}/L) T_a \quad , \\ FP(A) \exp(2\pi i \mathbf{k} \cdot \mathbf{x}/L) E_\alpha &= \frac{2\pi}{L^2} \mathbf{k} \cdot (2\pi \mathbf{k} + \mathbf{c}_\alpha) \exp(2\pi i \mathbf{k} \cdot \mathbf{x}/L) E_\alpha \quad , \end{aligned} \tag{28}$$

where $\mathbf{k} \neq \mathbf{0}$ and $\mathbf{f}^I \cdot \mathbf{k} \in \mathbb{Z}$. Therefore $\rho(c_i^a T_a / L) = 0$ if and only if there is an α and $\mathbf{k} \neq \mathbf{0}$, such that $\mathbf{k} \cdot \mathbf{c}_\alpha = -2\pi \mathbf{k} \cdot \mathbf{k}$. It defines a hyperplane of codimension 1, such that the Gribov horizon truely deserves its name: The origin is surrounded by a closed surface on which the Faddeev-Popov determinant vanishes, whereas in the interior all the eigenvalues of the operator $FP(A)$ are positive. This surface is specified most easily in terms of $c_I^{(i)}$ and amounts to $|c_I^{(i)}| \leq 2\pi$, with equality for at least one of the pairs $(I, (i))$. In particular the Z_G^3 conjugates of the origin are on the Gribov horizon. We have illustrate by the fat hexagon in figure 1 the position of the Gribov horizon.

Note that we analyzed the gauge fixing in the time independent three dimensional context. That is how it occurs in the Hamiltonian formulation [18], where $|\rho(A)|$ forms the integration measure for the gauge fixed potentials. In the derivation of the effective Hamiltonian [1] this measure was absorbed in the wave functional by the rescaling $\Psi(A) \to \sqrt{|\rho(A)|}\Psi(A)$. As a consequence, the wave functional will *vanish* at the Gribov horizon. Since the interior of the Gribov horizon, after dividing out the action of the constant gauge transformations, will cover the configuration space precisely once (to be honest we only proved this for the vacuum valley), we could define the theory on this space and impose vanishing boundary conditions on the horizon. However, this is not quite a practical approach. Apart from some technical reasons, having to do with the adiabatic approximation, the main objection is that we have destroyed the powerful symmetry of the wave functional under homotopically non trivial gauge transformations. In particular, the Z_G^3 conjugate vacua are not quite recognizable as equivalent vacua, around which one has as much right to expand as around zero.

It should be clear by now that the Gribov horizon can be considered merely as a coordinate singularity, in spirit rather reminiscent to the problems at the origin for a radially symmetric wave function. The Gribov horizon therefore arises due to the topological non trivial nature of the configuration space and the most obvious resolution of the problem of dealing with Gribov horizons, which, at the same time will restore the symmetries so dear to us, is to cover configuration space by coordinate patches. We choose as many patches as there are homotopy types of gauge transformations, each centered at the appropriate vacuum (the gauge copy of $\mathbf{A}=0$ under the appropriate homotopically non trivial gauge transformation). Thus, in our case we have patches $V_{P,\mathbf{\Theta}}$ centered at $A_j=-ig(P,\mathbf{\Theta})\partial_j g^{-1}(P,\mathbf{\Theta})$, where

$$g(P,\mathbf{\Theta})=g_1^P g_{\mathbf{\Theta}} \quad , \quad P\in\mathbb{Z} \quad , \quad \Theta_I\in\tilde{\Lambda}_w/\tilde{\Lambda}_r \quad . \tag{29}$$

The transition functions are nothing but the gauge transformations:

$$[g(P,\mathbf{\Theta})g^{-1}(\tilde{P},\tilde{\mathbf{\Theta}})]:V_{\tilde{P},\tilde{\mathbf{\Theta}}}\to V_{P,\mathbf{\Theta}} \quad . \tag{30}$$

In each patch we will define the Coulomb gauge as $\partial_i D_i([g^{-1}(P,\mathbf{\Theta})]A)=0$, for $A\in V_{P,\mathbf{\Theta}}$ and we will denote by $\rho_{P,\mathbf{\Theta}}(A)$ the corresponding Faddeev-Popov determinant. The Hamiltonian on $V_{P,\mathbf{\Theta}}$ is then the conjugate under $[g(P,\mathbf{\Theta})]$ of the Hamiltonian on $V_{0,\mathbf{0}}$. By arranging things this way the gauge condition is automatically preserved under the action of the transition functions, and the (rescaled) wave functional transforms as:

$$\begin{aligned}
&[g(P,\mathbf{\Theta})g^{-1}(\tilde{P},\tilde{\mathbf{\Theta}})]\Psi^{\theta,\mathbf{e}}_{\tilde{P},\tilde{\mathbf{\Theta}}}(A')/\sqrt{\rho_{\tilde{P},\tilde{\mathbf{\Theta}}}(A')}=\\
&\qquad\qquad \exp(i\theta(P-\tilde{P}))\exp(2\pi i\mathbf{e}\cdot(\mathbf{\Theta}-\tilde{\mathbf{\Theta}}))\Psi^{\theta,\mathbf{e}}_{P,\mathbf{\Theta}}(A)/\sqrt{\rho_{P,\mathbf{\Theta}}(A)} \quad ,\\
&\qquad A'=[g(P,\mathbf{\Theta})g^{-1}(\tilde{P},\tilde{\mathbf{\Theta}})]A \quad .
\end{aligned} \tag{31}$$

We wish to emphasize that the above is a precise description of how to define the Hamiltonian formulation, if one can prove that the union of the coordinate patches will cover the configuration space ($V_{P,\theta}$ is defined as the open set centered at $A_j=-ig(P,\mathbf{\Theta})\partial_j g^{-1}(P,\mathbf{\Theta})$ and bounded by the the Gribov horizon). We proved it for the vacuum valley, which is sufficient for our present purposes.

We now describe the process of reducing the problem of constructing the wave functional to a problem with a finite number of degrees of freedom. We would anticipate the zero-momentum modes to serve that purpose, however, what is zero-momentum expanding around the origin is no longer zero-momentum when expanding around the equivalent vacua. It is important though, that what are vacuum valley modes when expanding around the origin remain vacuum valley modes expanding around the equivalent vacua. If we could restrict the wave functional to the vacuum valley we would get simple Bloch type wave functions in a periodic potential. What stops us are the problems near the orbifold singularities, which forced us to include modes transverse to the vacuum valley. These modes, however, differ from patch to patch.

The resolution lies in understanding how one goes beyond the simple lowest order Born-Oppenheimer approximation. Consider for the sake of simplicity a Hamiltonian depending on two coordinates (x,y), of which x corresponds to the vacuum valley coordinate and y to the transverse coordinate.

$$H=-\frac{g^2}{2}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+\frac{1}{2g^2}(x^2-1)^2y^2 \quad . \tag{32}$$

For a slightly more realistic model see ref. [15], but this Hamiltonian [20] mimics enough of the features of the full problem to illustrate the procedure. We write the wave function as

$$\Psi(x,y)=\sum_{n=0}^{\infty}\phi_n(x)\chi^n_{[x]}(y) \quad , \tag{33}$$

where $\chi^n_{[x]}$ are the eigenfunctions for the transverse fluctuations (in this simple case the harmonic oscillator with frequency $\omega(x) = |x^2 - 1|$), defined by freezing the slow fluctuations (i.e. by dropping the terms proportional to $\partial/\partial x$). If $V_n(x)$ are the eigenvalues (here $(2n+1)\omega(x)/2$), we can rewrite the Hamiltonian equation in an infinite component form:

$$\Bigl[-\frac{g^2}{2}\Bigl(\delta_{nm}\frac{\partial}{\partial x} + A_{nm}(x)\Bigr)^2 + \delta_{nm}V_m(x)\Bigr]\phi_m(x) = E\phi_n(x) \quad , \tag{34}$$

$$A_{nm}(x) = <\chi^n_{[x]}|\frac{\partial}{\partial x}|\chi^m_{[x]}> \quad . \tag{35}$$

One can now show that, although near the conic singularities of the effective vacuum valley potential $V_o(x) = |x^2 - 1|/2$ the wave function has many components of ϕ_n non-negligible, that sufficiently far from these singularities (see [15,20] for the precise definition of this) ϕ_o will strongly dominate.

This principle can be applied to the gauge theory case. Furthermore, we can derive the decomposition of the wave function in terms of vacuum valley coordinates just as well starting from the effective Hamiltonian for the zero-momentum modes, eq. (26). This will turn out to be advantageous, because the symmetries imposed can be reformulated in terms of boundary conditions within one coordinate patch. Again there is a simple example to illustrate this principle for a Z_2 symmetry. Take the double anharmonic oscillator $H = (-g^2\partial^2/\partial x^2 + (x^2-1)^2/g^2)/2$, which has the symmetry $x \to -x$. We could define the theory on two coordinate patches, say $V_1 = \{x > -1/2\}$ and $V_2 = \{x < 1/2\}$, with a unit transition function. Wave functions will split into even (Ψ_e) and odd (Ψ_o) wave functions, which can be specified by $\partial\Psi_e(0)/\partial x = 0$ and $\Psi_o(0) = 0$. Thus we can define the theory in terms of wave functions on $\hat{V}_1 = \{x \geq 0\}$ (or $\hat{V}_2 = \{x \leq 0\}$) with a boundary condition on Ψ at the boundary of $\hat{V}_1$ (respectively $\hat{V}_2$).

Let us first discuss the case of SU(2). We introduce gauge invariant parameters (for simplicity we shall restrict ourselves to orthogonal lattices Σ, for which $\mathbf{f}^I \cdot \mathbf{f}^J = 0$ when $I \neq J$)

$$r_I^2 = \sum_{a=1}^{3} c_I^a c_I^a = 2\mathrm{Tr}(c_I^2) \quad , \tag{36}$$

which are associated with the quadratic Casimir of SU(2). The remaining coordinates are angular, parametrizing the unit vectors c_I/r_I, whereas the vacuum valley is given by those vectors that are (anti-)parallel. We therefore choose the vacuum valley parameters to be x_I, with $r_I = |x_I|$, whereas a simultaneous change of the sign of these coordinates is equivalent to a gauge transformation (for SU(2) the Weyl group is indeed isomorphic to Z_2). Furthermore, next to the gauge symmetry one has the invariance group of the lattice Σ. For a cubic lattice this would be the cubic group $O(3, \mathbb{Z})$, for which we discussed the representations in the introduction. However, any orthogonal lattice has the reflection symmetries $P_I : c_J^a \to \pm c_J^a$, with a positive sign for $I \neq J$ and a minus sign for $I = J$. The product of these reflection symmetries is the parity $P = P_1P_2P_3$. These symmetries descent to the vacuum valley, but the parity transformation on the vacuum valley is equivalent to a gauge transformation (since $\tau_1\tau_3\tau_1 = -\tau_3$, with τ_k the Pauli spin matrices). Hence, negative parity wave functionals cannot be reduced to the vacuum valley [21] ($\phi_o = 0$ in the equivalent of eq. (33)). For a more detailed discussion and a possible remedy we refer to ref. [5]. We thus restrict to the case of positive parity.

The fundamental domain of the vacuum valley in terms of which we will phrase the boundary conditions is $-\pi \leq x_I \leq \pi$, which has the property that its boundary is fixed under a Z_2 subgroup of the symmetries. By choosing $\mathbf{\Theta}$ appropriately we can arrange things such that

$$[g_{\mathbf{\Theta}}] \circ P_I : x_I \to 2\pi - x_I \quad , \tag{37}$$

leaving x_J for $I \neq J$ untouched. We therefore get as boundary conditions (p_I is the eigenvalue of the reflection symmetry P_I):

$$\begin{aligned} \phi_o(\mathbf{x}) &= p_I(-1)^{e_I}\phi_o(\mathbf{x}) \quad , \\ \frac{\partial}{\partial x_I}\phi_o(\mathbf{x}) &= -p_I(-1)^{e_I}\frac{\partial}{\partial x_I}\phi_o(\mathbf{x}) \quad , \quad \text{at} \quad x_I = \pi \quad . \end{aligned} \tag{38}$$

These boundary conditions, since they are formulated in terms of gauge invariant coordinates, can be rephrased in terms of the wave function for the zero-momentum modes (see eq. (26)):

$$\begin{aligned} \frac{\partial}{\partial r_I}(r_I\Psi(c)) &= -p_I(-1)^{e_I}\frac{\partial}{\partial r_I}(r_I\Psi(c)) \quad , \\ \Psi(c) &= p_I(-1)^{e_I}\Psi(c) \quad , \quad \text{at} \quad r_I = \pi \quad , \end{aligned} \tag{39}$$

where the factor r_I in front of the wave function arises due to a Jacobian factor, which is a remnant of the Faddeev-Popov determinant $\rho(A)$. The reason we can infer that this is all that is left of $\rho(A)$, when going to the effective theory for the zero-momentum modes, is that with these boundary conditions we can show [4] that the analogue of eq. (34) gives a continuous energy density across the various coordinate patches in the vacuum valley. We will come back to this point in the context of arbitrary groups. The approximation that the wave function is predominantly in the ground state for the transverse fluctuations, is called the adiabatic approximation. The boundary conditions of eq. (39) are equivalent to the condition that the energy density *as derived from the effective Hamiltonian for the zero-momentum modes* is continuous. The adiabatic approximation is necessary only to argue that this gives a good approximation to the full problem. For SU(2) the validity of this adiabatic approximation was tested extensively [4].

Since we have now reduced the problem to that of an ordinary Hamiltonian quantum mechanics problem, we proceed by choosing a suitable basis of wave functions that all satisfy the boundary conditions and are gauge invariant. We can then diagonalize the Hamiltonian in a Rayleigh-Ritz analysis, from which we could derive both upper and lower bounds on the energy eigenvalues [4]. The particular basis of these wave functions is

$$\Psi_{\ell,\mathbf{n},\mathbf{e}}(c) = \sum_{\mathbf{m}} W(\ell,\mathbf{m}) \prod_{I=1}^{3} \chi^{(e_I)}_{n_I,\ell_I}(r_I) Y_{\ell_I,m_I}(\theta_I,\varphi_I) \quad , \tag{40}$$

where $W(\ell,\mathbf{m})$ are the Clebsch-Gordon coefficients for the product of three SO(3) multiplets projected onto a singlet, so as to guarantee the gauge invariance. The radial wave functions $\chi^{(e)}_{n,\ell}(r)$ satisfy the boundary conditions

$$\begin{aligned} \frac{d}{dr}(r\chi^{(e)}_{n,\ell}(r)) &= -(-1)^{\ell+e}\frac{d}{dr}(r\chi^{(e)}_{n,\ell}(r)) \quad , \\ \chi^{(e)}_{n,\ell}(r) &= (-1)^{\ell+e}\chi^{(e)}_{n,\ell}(r) \quad , \quad \text{at} \quad r = \pi \quad . \end{aligned} \tag{41}$$

A convenient choice for these radial wave functions are the spherical Bessel functions $j_\ell(kr)$, where k is determined from the boundary conditions and is labelled by e and n (the number of nodes).

Now we will review the recent analysis of Vohwinkel [3], who was able to generalize the above construction to the gauge group SU(3), which is more realistic if we ultimately wish to understand QCD. Since SU(2) is not quite generic such a generalization is not entirely straightforward, particularly with respect to formulating the appropriate boundary conditions. We introduce again gauge invariant parameters associated with the Casimir invariants (contraction of the c_I^a with the invariant tensors of the gauge group). For SU(3) this leads to the coordinates

$$r_I^2=\sum_{a=1}^{8}c_I^ac_I^a=2\mathrm{Tr}(c_I^2)\quad,\quad s_I=\sum_{a,b,e=1}^{8}c_I^ac_I^bc_I^ed_{abe}=4\mathrm{Tr}(c_I^3)\quad,\tag{42}$$

These can be easily reformulated in terms of vacuum valley coordinates by substituting an abelian configuration

$$c_I=x_IT_1+y_IT_2\quad,\tag{43}$$

where in terms of the standard Gell-Mann matrices λ_a we have chosen $T_1=\lambda_8/2=\mathrm{diag}(1,1,-2)/\sqrt{12}$ and $T_2=\lambda_3/2=\mathrm{diag}(1,-1,0)/2$. We find for these configurations

$$r_I^2=x_I^2+y_I^2\quad,\quad s_I=\sqrt{3}x_I(y_I^2-x_I^2/3)\quad,\tag{44}$$

which can be solved as

$$x_I=r_I\sin\varphi_I\quad,\quad y_I=r_I\cos\varphi_I\quad,\quad \sin(3\varphi_I)=\frac{s_I}{\sqrt{3}r_I^3}\quad.\tag{45}$$

Note that (x_I,y_I) can be identified with (c_I^1,c_I^2), introduced when discussing the vacuum valley, but in order to stress their gauge invariance (modulo the action of the Weyl group), we prefer the present notation to avoid possible confusion. As for SU(2) we have to extend these variables to the whole vacuum valley, since (r_I,s_I) will only parametrize the Weyl chamber [17], indicated in figure 1 by the shaded area. Any point in the dual weight lattice can be obtained by applying a Weyl reflection to a point in this Weyl chamber. We have to remember though, that the gauge transformations act on the vacuum valley as a simultaneous Weyl reflection on each of the three directions. Hence we will take the coordinates (x_I,y_I) to parametrize the vacuum valley, but we will choose the basis of wave functions by "radial" wave functions χ depending on (r_I,s_I) only and "angular" wave functions Y that form irreducible representations of the adjoint representation of SU(3) and that are independent of the coordinates (r_I,s_I). Thus (x_I,y_I) is labelled by (r_I,s_I) *and* the Weyl reflection that is necessary to bring (x_I,y_I) to the fundamental Weyl chamber. This is in complete analogy with SU(2), if we realize that x_I is labelled by r_I and the reflection P_I that might be required to relate x_I to r_I (the fundamental Weyl chamber for SU(2) is given by $x_I\geq 0$). Note that under the reflection symmetries, r_I behaves as an even and s_I as an odd function, so that total parity, when restrict to the vacuum valley, is not equivalent to a gauge transformation. Therefore, for SU(3) we can consider negative parity wave functionals [3].

However, since SU(3) has complex representations, there is an enhanced symmetry of charge conjugation. It inverts the sign of the abelian generators, but preserves the structure constants (thereby requiring specific transformation rules for E_α, which are not very relevant for the present discussion). Thus, on the vacuum valley, charge conjugation acts as a parity transformation. In particular the combination of parity and charge conjugation, CP, will act trivially on the vacuum valley. This means that we can *not* restrict wave functionals with $CP=-1$ to the vacuum valley.

It is interesting to observe that, when parity (or charge conjugation) restricted to the vacuum valley is not equivalent to a gauge transformation, this implies that there is an automorphism of the root system which is not a Weyl transformation. It is indeed well known [17] that the group of automorphisms of the root system is a semidirect product $\Gamma\mathcal{W}$, where Γ is the group that leaves fixed a given basis for the root system. For SU(2) one has $\Gamma=1$, essentially because there is only one root, confirming that parity is equivalent to a gauge transformation. On the other hand for SU(N) with $N>2$, $\Gamma=Z_2$. For SU(3) this corresponds to reflection in the line through δ (see figure 1), which up to a conjugation with a Weyl reflection corresponds to $x_I\to -x_I$. It is interesting to note that there is only one Lie-algebra for which $\Gamma\neq 1$ and $\Gamma\neq Z_2$. This is the algebra belonging to the root system D_4. The corresponding simply connected simple gauge group is Spin(8) which is the covering group of SO(8). For this case $\Gamma=\mathcal{S}_3$, which is related to the famous triality of SO(8).

Coming back to SU(3), we will now formulate the boundary conditions on the "radial" part of the wave function [3]. As for SU(2) these are determined by decomposing the wave function as in eq. (33), formulating eq. (34) such that the Hamiltonian density can be continuously extended over the whole vacuum valley. One can convince oneself easily from the fact that this implies that the kinetic term for the vacuum valley coordinates should be formulated with respect to a flat metric. This is achieved by absorbing the square root of the Jacobian, associated to the change of variables from c_I^a to (x_I,y_I,Ω_I) (where Ω_I stands for the collection of angular variables), in the wave function. As mentioned before, this Jacobian is a remnant of the Faddeev-Popov determinant and it factorizes as the product of Jacobians, $\hat\rho_I=J_I^2$, for each direction. The functions $\hat\rho_I$ are gauge invariant and henceforth we will suppress the index I, since we can formulate the boundary conditions independently for each direction. The function J is assumed to be analytic in the vacuum valley coordinates and the gauge invariance of its absolute value therefore implies that $\mathrm{sign}(J)$ is the completely antisymmetric singlet representation of the Weyl group. One easily verifies this from their explicit forms:

$$\mathrm{SU(2)}:\quad J=x\quad,\quad \mathrm{SU(3)}:\quad J=y(y^2-3x^2)\quad. \tag{46}$$

The formulation of the boundary conditions is equivalent to formulating how one patches together the function $\hat\chi$ defined by

$$\hat\chi(X)\hat Y(\Omega)=\sqrt{\hat\rho(R)}\chi(R)Y(\Omega)\quad. \tag{47}$$

We suppressed the labels and X stands for the collection of vacuum valley coordinates, which are Weyl equivalent to the gauge invariant "radial" coordinates R. The "angular" coordinates are denoted by Ω and $Y(\Omega)$ is in an irreducible representation of the group of gauge rotations. The function $\hat Y(\Omega)$ is obtained from $Y(\Omega)$ by eliminating the dependence under the Weyl reflections

$$Y(\Omega)=Z(w)\hat Y(\Omega)\quad, \tag{48}$$

where w denotes a Weyl chamber. Note that $Z(w)$ is a representation of the Weyl group, which can be decomposed in its irreducible representations $Z_{irr}(w)$. For SU(2) the Weyl group is Z_2 and has two irreducible representations, the trivial and the completely antisymmetric representation, which are both singlets. Explicitly one has $\hat{Y}_{\ell m}(\Omega) = \mathrm{sign}^{\ell}(x) Y_{\ell m}(\Omega)$. For SU(3) the Weyl group is $\mathcal{S}_3$, which has three irreducible representations: The trivial representation A_1, the antisymmetric representation A_2 ($= \mathrm{sign}(J)$), and the doublet E. Since $\mathcal{S}_3$ is a normal subgroup of the cubic group, we can use the notation for the cubic group, discussed in the introduction. It is instructive to specify these irreducible representation in a little more detail to illustrate something that will allow generalization to arbitrary groups. The Weyl group $\mathcal{S}_3$ is generated by the rotations $\hat{R}$ over $2\pi/3$ and the reflections $\hat{P}$ in the x-axis. The irreducible representations are described by specifying the action of these transformations on a state vector Z_1.

$$\begin{array}{rl} A_1: & \hat{R}Z_1 = \quad \hat{P}Z_1 = Z_1 \quad , \\ A_2: & \hat{R}Z_1 = -\hat{P}Z_1 = Z_1 \quad , \\ E: & \hat{R}Z_1 = \quad \hat{P}Z_1 = Z_2 \quad , \quad \hat{R}Z_2 = \hat{P}Z_2 = -Z_1 - Z_2 \quad . \end{array} \tag{49}$$

If we denote by Z_{irr} the state vector Z_1 defining each of the irreducible representations, one easily finds that $Z_{A_1} = 1$, $Z_{A_2} = \mathrm{sign}(J)$ and Z_E is an arbitrary linear combination of $Z_E^{\pm}$, which we have indicated in figure 2. Note that each Z has unit modulus, such that they are mutually orthonormal and that Z_E^{+} is the complex conjugate of Z_E^{-}, which are interchanged under reflection in the y-axis.

Figure 2: The hexagons correspond to the coordinate patch indicated in figure 1 by the dashed hexagon. Each sector is a Weyl chamber in which we give the value of the state vector Z_E^{+}, where $z = \exp(2\pi i/3)$. (The state vector Z_E^{-} is obtained by a reflection in the vertical.) The matching factors $z^{(\mathcal{I})}$ in patching the function $\hat{\chi}$ to a globally defined function ϕ are such that $z^{(\mathcal{I})} Z_{irr}$ for a neighbouring patch has the same value in the sectors that touch each other at the boundaries.

We now split $\hat{\chi}$ according to the irreducible representations of the Weyl group as a linear combination of $\hat{\chi}_{A_1}$, $\hat{\chi}_{A_2}$ and $\hat{\chi}_E^{\pm}$, such that in each patch $(\mathcal{I})$ we have

$$\hat{\chi}_{irr}^{(\mathcal{I})} = J^{(\mathcal{I})}(X)\chi(R)\mathrm{sign}(J^{(\mathcal{I})}(X))Z_{irr}^{(\mathcal{I})}(w) \quad . \tag{50}$$

The patches are defined such that their boundaries are invariant (as a set) under translations over $2\pi\tilde{\Lambda}_w$ and under Weyl transformations. In figure 1 the dashed hexagon indicates the boundary of the patch centered at the origin, the other patches are obtained by translation over $2\pi\tilde{\Lambda}_w$. Furthermore, $\hat{\chi}_{irr}$ has to vanish along the boundaries of the Weyl chambers, which we enforce by demanding it to be odd with respect to all Weyl reflections, that generate the Weyl group. We define $\hat{\chi}_{irr}$ globally by introducing the function ϕ_{irr}, which coincides with $\hat{\chi}_{irr}$ in one of the patches (say the one centered at the origin), but differs by a factor $z^{(\mathcal{I})}$ (in the center of the gauge group) from $\hat{\chi}^{(\mathcal{I})}$, due to the matching of $\mathrm{sign}(J^{(\mathcal{I})})Z_{irr}^{(\mathcal{I})}$ to $\mathrm{sign}(J)Z_{irr}$ as defined in the patch centered at the the origin. We illustrated this for Z_E^+ in figure 2. Thus we find the following properties which will completely specify the boundary conditions of ϕ_{irr} (together with ϕ_{irr} being odd under Weyl reflections)

$$\begin{aligned}
\mathrm{SU}(2): \quad & \phi_{A_k}(x+2\pi) = \exp(\frac{2\pi i(e+k)}{2})\phi_{A_k}(x) \quad , \\
\mathrm{SU}(3): \quad & \phi_{A_k}(x+\frac{4\pi}{\sqrt{3}},y) = \exp(\frac{2\pi i e}{3})\phi_{A_k}(x,y) \quad , \\
& \phi_E^{\pm}(x+\frac{4\pi}{\sqrt{3}},y) = \exp(\frac{2\pi i(e\pm 1)}{3})\phi_E^{\pm}(x,y) \quad ,
\end{aligned} \tag{51}$$

where $k=1$ or $k=2$. For SU(2), A_1 corresponds with even ℓ, whereas A_2 corresponds with odd ℓ and the extra factor $2/\sqrt{3}$ for SU(3) arises from the fact that $\exp(2\pi i\lambda_8/\sqrt{3})$ generates Z_3 (as opposed to $\exp(\pi i\tau_3)$, which generates Z_2) or from the explicit expression of the dual weight vector $\tilde{\lambda}_{(1)} = (2/\sqrt{3},0)$. We can now easily find a complete set of wave functions with respect to which we can diagonalize the Hamiltonian, by letting $Y(\Omega)$ run over all inequivalent irreducible representations of the group of gauge rotations and by letting ϕ run over all possible eigen functions of the Laplacian

$$\sqrt{\hat{\rho}}\frac{\partial^2(\phi/\sqrt{\hat{\rho}})}{\partial c^a\partial c^a} = (\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\cdots)\phi = \mu\phi \quad . \tag{52}$$

It is easy to explicitly write down these eigen functions for SU(2) and SU(3):

$$\begin{aligned}
SU(2): \quad \phi(x) =& \sin(nx/2) \quad , \\
SU(3): \quad \phi(x,y) =& \sin(n_2 y/2)\exp(in_1\sqrt{3}x/2) + \\
& \sin(n_2(\sqrt{3}x-y)/4)\exp(in_1\sqrt{3}(\sqrt{3}x+y)/2) - \\
& \sin(n_2(\sqrt{3}x+y)/4)\exp(in_1\sqrt{3}(\sqrt{3}x-y)/2) \quad .
\end{aligned} \tag{53}$$

We will leave it to the reader to associate with each electric flux e and irreducible representation of the Weyl group the appropriate set of integers n or (n_1,n_2).

It is crucial to note now that both for SU(2) and SU(3) ϕ is invariant with respect to translations over $2\pi\tilde{\Lambda}_r$

$$\phi(X+2\pi\tilde{\Theta}) = \phi(X) \quad , \quad \tilde{\Theta}\in\tilde{\Lambda}_r \quad , \tag{54}$$

which is a consequence of the invariance under homotopically trivial gauge transformations and therefore generalizes to arbitrary groups. Thus, there is no fundamental obstacle to generalize the construction of the effective theory, restricted to one coordinate patch, to arbitrary gauge groups. Although all group theoretic arguments remain unaltered, going to bigger groups requires some hard work for (i) constructing the irreducible representations of the group of gauge rotations, (ii) decomposing these in the appropriate irreducible representation of the Weyl group and finally (iii) calculating all the matrix elements of the Hamiltonian with respect to the chosen basis of wave functions which, constructed for each direction separately are projected by the use of Clebsch-Gordan coefficients onto the singlet representation of the gauge group, to enforce invariance of the wave function under constant gauge transformations.

In the analysis there are furthermore three things that are different from SU(2) for arbitrary gauge groups. For SU(2), the matrix elements could be deduced from the matrix elements for each direction separately, which is possible because the effective Hamiltonian for SU(2) is invariant under Weyl reflections in each direction separately (essentially because the Weyl reflections could be identified with the coordinate reflections). For larger groups this is no longer the case. For SU(3), Vohwinkel [3] showed that this problem only occurs for terms of sixth (and higher) order in the fields and approximated the calculation of the matrix elements by replacing this term by its average over the Weyl reflections in each direction separately. This is expected to lead to only small errors. Also different from SU(2) is that there are now singularities (the boundaries of the Weyl chambers) that connect the origin to the points conjugate under the action of $Z_{\tilde{G}}$ (see figure 1), which is where breakdown of the adiabatic approximation is expected. Nevertheless, there are good reasons to expect that also for SU(3) the adiabatic approximation will be reasonable, but an analysis similar to what was done for SU(2) [4] is much harder for larger groups and it still will not allow one to easily estimate the error introduced by the fact that not all of the wave function is in the ground state for the transverse fluctuations (which has the possibility of leading to a coupling of the modes which were integrated out in one patch, but not in the neighbouring patch). Finally, since for larger groups there will be more quantum numbers to classify the basis vectors and since the radial functions alone will diagonalize the Laplacian (not incorporating that for large angular momenta the wave functions will be suppressed near the origin) one might have to take a substantially larger number of basis vectors to get an accurate value for the energies. Estimating lower bounds, as was done for SU(2) [4] will, however, become unpractical since it will involve calculating the matrix elements of higher powers of the Hamiltonian. Nevertheless, especially for larger effective couplings, the wave functions will spread out over the vacuum valley and one expects that the basis vectors with low angular momenta will strongly dominate in the larger volume range.

RESULTS AND CONCLUSIONS

We will now discuss some of the results and compare with the Monte Carlo calculations. For reviews of the status of pure gauge theory, we refer to [10,22,23]. The comparison between the Monte Carlo and the analytic calculations is made in terms of scale invariant quantities, for which we consider mass ratios and the parameters $z = mL$, i.e. the size of the torus in units of the correlation length determined by the mass m. Another quantity that played an important role, is the energy of electric flux as defined in the introduction. The string tension κ is by definition the energy of electric flux divided by the size L of the torus. Its square root scales as a mass; it and the other masses are expected to tend to a constant when the volume goes to infinity. The analytic calculations have now been applied to SU(2) in cubic and asymmetric volumes. The asymmetric volume, described by an orthogonal lattice Σ where one of the vectors $\mathbf{f}^I$ differs in length from the other two, has been used to estimate the deconfining temperature [10,24]. Most recently Vohwinkel generalized both the cubic and asymmetric volume calculations to SU(3) [3].

We will here only illustrate the results for SU(2) and cubic volumes. With the increasing accuracy of the Monte Carlo calculations, especially using the "fuzzing" technique to build operators that couple well to the various glueball states [11], it became clear that there are systematic differences between the Monte Carlo and the analytic calculations. In an attempt to understand those differences we derived the effective Hamiltonian for the zero-momentum modes from Wilson's lattice action, thereby explicitly including the lattice artifacts [12,25]. It is important to observe that gauge transformations that are only periodic up to an element in the center of the gauge group also exist on a finite lattice, and this leads to an exact symmetry of the transfer matrix. Thus, electric flux is a proper quantum number for a finite lattice, unlike θ which is only defined in the continuum limit. It follows that the complete analysis of this paper for the vacuum valley carries over to the lattice. In figure 3 we have compared the analytic calculations for the ratio of the square root of the string tension with the glueball mass in the E^+ representation of the cubic group as a function of the parameter $z_{E^+} = m_{E^+}L$, for lattices of spatial size 4^3, 6^3 and ∞^3 (which is equivalent to the continuum). The dashed line includes a two-loop correction to the effective potential [12] and the Monte Carlo data are from refs. [10,11].

Figure 3: (From ref. [12]) Comparison of the Monte Carlo results (a from ref. [10] and b from ref. [11]) with the analytic results (the dashed curve includes a two-loop correction). Horizontally is plotted $z_{E^+} = m_{E^+}L$ and vertically $\sqrt{\kappa}/m_{E^+}$, where κ is the string tension.

In figure 4 we made a comparison for the mass ratio $m_{A_1^+}/m_{E^+}$ (for this case the two-loop correction turns out to be negligible). The Monte Carlo data here were based on those of ref. [11] (the scalar glueball mass needs a very "fuzzy" operator to avoid the mixing with the higher lying states observed in ref. [10]). The excellent agreement speaks for itself.

As we already mentioned, the calculations can not be expected to be valid all the way to an infinite volume, because for sufficiently strong coupling the θ dependence is expected no longer to be negligible. At the more technical level, in the derivation of the effective theory it is assumed that we can separate the zero-momentum from the non-zero momentum modes. Naively this approximation can be expected to break down for masses that are of the order of the energy of the lowest non-zero momentum mode, which means that the method is expected to break down for $z = mL > 2\pi$, which is compatible with the general consensus that the effective Hamiltonian cannot describe the region $z_{E^+} > 5$, which in physical units for a glueball mass of about 1.3 GeV corresponds with distances of 0.7 fermi. Nevertheless, we hope that the breakdown of the approximation fits naturally in our scheme and is caused by the fact that the energies become of the order of the height of the potential barrier, which separates two components of the vacuum. From there on the θ dependence can no longer be neglected. This is confirmed in SU(3) by recent results for the topological susceptibility (which is the second derivative of the vacuum energy density with respect to θ), which as a function of the size of the torus is seen to suddenly increase from zero to a large value around $z_{A_1^{++}} = 5.9$ [26]. For SU(2) there is also some indirect evidence for the sudden onset of θ dependence [27].

Figure 4: (From ref. [25]) Comparison of the Monte Carlo [11] with the analytic results for the mass ratio $m_{A_1^+}/m_{E^+}$. Horizontally is plotted $z_{E^+} = m_{E^+}L$. The circles are from 4^3N_o and the squares from 6^3N_o lattices. Here the two-loop correction is negligible.

In ref. [28] was speculated on a mechanism of domain formation, driven by the instanton dynamics and it is interesting that the domain walls in such a picture are of a magnetic nature, and that magnetic monopoles (as gauge dislocations, occurring in "matching" the domains) will naturally occur. Since we have not yet been able to develop these ideas in a more quantitative context, we will say no more except for pointing to an additional piece of circumstantial evidence [22], namely that the multiplicity weighted average of the T_2^+ and E^+ mass in the intermediate volume agrees well with the large volume result for the mass of the 2^+ (tensor) glueball, determined from lattice Monte Carlo calculation. Since the 2^+ representation of the full rotation group decomposes in to the E^+ and T_2^+ representation of the cubic group, this is what one can expect in a domain type vacuum, when rotational invariance is restored by averaging over all the orientations of the domains. From this point of view it would be very desirable if the error in the large volume Monte Carlo determination of the tensor glueball mass could be brought down considerably.

In conclusion, we have illustrated that there is a very good agreement with the Monte Carlo results in intermediate volumes, where θ dependence can be neglected, so that we can be confident to have understood how to go beyond the Gribov horizons belonging to the topologically non-trivial gauge transformations associated with $H_1(T^3, Z_G)$. The challenge for the future is to also enlarge our horizon beyond the Gribov problems associated with $\pi_3(G)$, which would create the exciting possibility of calculating the spectrum as a function of θ.

ACKNOWLEDGEMENTS

My gratitude goes to Poul Damgaard for inviting me to this workshop in the splendid surroundings of Cargèse and for the financial support. I thank Henrik Flyvbjerg, Tony Kennedy, Bo Söderberg and the other participants for stimulating discussions and Mees de Roo for a discussion on triality in SO(8). Finally it is a pleasure to acknowledge illuminating discussions with Claus Vohwinkel.

REFERENCES

[1] M. Lüscher, Nucl. Phys. B219(1983)233; M. Lüscher and G. Münster, Nucl. Phys. B232(1984)445

[2] G. 't Hooft, Nucl. Phys. B153(1979)141

[3] C. Vohwinkel, Analytic Progress Towards the Mass-Spectrum and Deconfining Temperature in SU(3) Gauge Theory, Tallahasse preprint, September 1989, to appear in Phys. Rev. Lett.; Ph.D. thesis, Tallahassee, September 1989

[4] J. Koller and P. van Baal, Nucl. Phys. B302(1988)1

[5] P. van Baal, Acta Phys. Pol. B20(1989)295

[6] V. Gribov, Nucl. Phys. B139(1978)1

[7] R. Jackiw, I. Muzinich and C. Rebbi, Phys. Rev. D17(1978)1576

[8] I. Singer, Comm. Math. Phys. 60(1978)7

[9] W. Nahm, in: IV Warsaw Symp. on Elem. Part. Phys., ed. Z. Ajduk, p.275 (Warsaw, 1981)

[10] B.A. Berg and A.H. Billoire, Phys. Rev. 40(1989)550; B.A. Berg and C. Vohwinkel, Phys. Rev. 40(1989)584

[11] C. Michael, G.A. Tickle and M.J. Teper, Phys. Lett. 207B(1988)313; C. Michael, Liverpool preprint LTH 230, May 1989

[12] P. van Baal, Phys. Lett. B224(1989)397

[13] P. van Baal, Nucl. Phys. B307(1988)274; in: Frontiers in Nonperturbative Field Theory, eds. Z. Horvath, e.a., World Scientific, Singapore, 1989, p.204

[14] J. Kripfganz and C. Michael, Nucl. Phys. B314(1989)25; C. Michael, Liverpool preprint LTH 230, May 1989

[15] P. van Baal and J. Koller, Ann. Phys.(N.Y.) 174(1987)299

[16] P. van Baal, Comm. Math. Phys. 85(1982)529

[17] J.E. Humphreys, Introduction to Lie Algbras and their Representations, Springer, New York, 1972; V.S. Varadarajan, Lie Groups, Lie Algebras and their Representations, Springer, New York, 1984

[18] N.M. Christ and T.D. Lee, Phys. Rev. D22(1980)939

[19] C. Bloch, Nucl. Phys. 6(1958)329

[20] P. van Baal and A. Auerbach, Nucl. Phys. B275[FS17](1986)93

[21] C. Vohwinkel, Phys. Lett. B213(1988)54

[22] P. van Baal and A. Kronfeld, Nucl. Phys. B(Proc. Suppl.)9(1989)227

[23] C. Michael, Pure Gauge: Glueballs, Potentials and the Vacuum, to appear in the proceedings of LAT'89, Capri, Italy, Liverpool preprint LTH 244, October 1989

[24] B. Berg, C. Vohwinkel and C. Korthals-Altes, Phys. Lett. B209(1988)319

[25] P. van Baal, Understanding SU(2) Lattice Artifacts in Intermediate Volumes, to appear in the proceedings of LAT'89, Capri, Italy, Utrecht preprint THU-89/29, October 1989

[26] J. Hoek, Topology, String Tension and Glueball Masses in Intermediate Volume SU(3) Lattice Gauge Theory, Rutherford preprint RAL-89-081, August 1989

[27] A.S. Kronfeld, M.L. Laursen, G. Schierholz, C. Schleiermacher and U.-J. Wiese, Nucl. Phys. B305(1988)661

[28] J. Koller and P. van Baal, Nucl. Phys. B(Proc. Suppl.)4(1988)47