Nuclear Physics B 385 (1992) 185–226
North-Holland

NUCLEAR PHYSICS B

The theta dependence beyond steepest descent

Pierre van Baal [1]

Institute for Theoretical Physics, Princetonplein 5, P.O.Box 80006, 3508 TA Utrecht, The Netherlands

N.D. Hari Dass [2]

Institute for Theoretical Physics, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

Received 3 February 1992
Accepted for publication 20 July 1992

This paper outlines the strategy for computing the θ-dependence in non-abelian gauge theories beyond a semiclassical or steepest descent approximation. It involves isolating the relevant degrees of freedom including the sphaleron configuration for tunnelling across a classical potential barrier. Two approaches are discussed in the context of spherical geometries. The first is based on a hamiltonian version of the streamline or valley equation. The second, which in our opinion is far more efficient, is based on implementing θ-dependence through appropriate boundary conditions in configuration space. In a good approximation these can be formulated at the level of 15 (+3 gauge) modes, that are degenerate to lowest order in perturbation theory, while keeping all other modes gaussian.

1. Introduction

Tunnelling through classical potential barriers has been an important source of information for non-perturbative behaviour in non-abelian gauge theories [1,2]. Recently this attracted much attention in the context of the electroweak sector [3] also. In most situations the effect of tunnelling is computed through a steepest descent or semiclassical approximation. This is reasonable as long as the energy of the states for which these non-perturbative contributions are to be computed is below the minimal barrier energy. The saddle point corresponding to this minimal barrier energy is what became known as a sphaleron [4]. One would, however, also wish to know the non-perturbative contributions if the energies are comparable to and higher than the sphaleron energies. In the electroweak sector this is done by extrapolating the results obtained from the steepest descent calculations, based on a perturbative expansion around the instantons (i.e. the classical solutions of the euclidean

[1] KNAW fellow.
[2] On leave of absence from the Institute of Mathematical Sciences, Madras, India.

equations of motion that connect the two "nearest" classical vacuum configurations). In QCD or non-abelian gauge theories similar steepest descent approximations are plagued by infrared divergences [1,5], due to the integration over the instanton scale parameter. Because of the interacting nature of the infrared modes it is most sensible to formulate the theory in a finite volume [6-9]. Studying the low-energy dynamics of the non-abelian gauge theories can now be performed perturbatively, as asymptotic freedom makes the effective coupling small for small volumes. What one discovers when one increases the volume, is that wave functionals start to spread out over configuration space and degeneracies due to a multiple classical vacuum structure will be lifted progressively. This line of approach has been particularly fruitful if we make the volume finite by imposing periodic boundary conditions [6,8]. This has a richer classical vacuum structure than for example in spherical compactification [7]. In the absence of fermions the θ-parameter is the relevant quantum number to connect the wave functionals in the various vacua, whereas in the torus geometry one has in addition electric flux quantum numbers [9] $e_i \in \mathbb{Z}_G$ ($\mathbb{Z}_G$ is the center of the gauge group G). As long as the volume (and thus the effective coupling constant) is so small that the energies of the low-lying states are well below all sphaleron energies, perturbation theory will be appropriate. (The sphaleron energy associated to e_i is quantum induced [8], a complication that is irrelevant for our present arguments). At increasing volume, energies will become comparable to the electric flux sphaleron energy but will remain small with respect to the θ-sphaleron energy. For the torus geometry this is at volumes between $(0.2\,\text{fm})^3$ and $(0.8\,\text{fm})^3$, the scale being set by the physical string tension. In this so called intermediate volume domain the low-lying spectrum was analytically computed [8] and for SU(2) agrees perfectly with the most accurate Monte Carlo calculations [10] (with statistical errors of approximately 2%). A similar analysis was performed for SU(3) [11]. The main idea was to isolate the degrees of freedom that include the relevant sphaleron configurations and to derive an effective hamiltonian for these 6 (+3 gauge) degrees of freedom. This effective hamiltonian deviates from the one derived for the perturbative analysis [6] only by imposing boundary conditions on the wave functions at the sphaleron configurations.

These computations essentially go beyond a semiclassical approximation and have clearly demonstrated that extrapolating the result obtained from such an approximation [12] to the domain where energies are comparable to the sphaleron energy is inappropriate. Of course one can always a posteriori match some parameters of the semiclassical approximation to get a better fit to the actual results, but the problem is that only very limited knowledge about the analytic structure in the coupling constant of the relevant amplitudes is available, making extrapolation unreliable to a large degree. Thus, the claims of large cross sections at high energies for $(B + L)$-violating processes in the electroweak sector [3] will remain controversial as long as they are based on the steepest descent approximation of instanton contributions [13]. The main emphasis of this paper is, however, towards the QCD applications, but we hope that our techniques will ultimately be applicable in the more complicated setting of the electroweak theory also.

The comparison of the analytic intermediate volume computations with the lattice Monte Carlo results demonstrate that the effective hamiltonian becomes unreliable at volumes bigger than $(0.8\,\mathrm{fm})^3$ and it is most natural to assume that this is due to energies becoming comparable to the θ-sphaleron energy. This was confirmed by a Monte Carlo investigation of the topological susceptibility [14] (the second derivative of the ground state energy with respect to θ), which showed a near abrupt onset above a volume of approximately $(0.8\,\mathrm{fm})^3$. It is therefore most natural to enlarge the number of degrees of freedom to also include the θ-sphaleron. This seemingly straightforward extension was severely hampered by a technical obstruction. Neither the sphaleron, nor the instantons are explicitly known for the torus geometry. There has been over the years various fruitless attempts to construct instantons on T^4 and one can actually prove [15] that for unit topological charge no regular instanton exists. It is important to stress, as this result has often been interpreted incorrectly, that this does *not* imply absence of regular instantons for the (spatial) torus geometry if the time direction is not compactified too. A similar situation exists for a simple one-dimensional double-well $H = -\frac{1}{2}\partial^2/\partial x^2 + \frac{1}{2}\lambda^2(x^2-1)^2$, whose instanton (or kink) equation, $\mathrm{d}x/\mathrm{d}t = \pm\lambda(x^2-1)$, has no solutions with $x(T) = -x(-T) = \pm 1$, unless T approaches infinity.

There are two important reasons to stick to the torus geometry. One is practical, as it allows for comparisons with lattice Monte Carlo results, something that should not be abandoned too easily. It allows one to test the approximations in the analytic approach, whereas for the lattice it is a test of lattice artifacts [12c]. The second, equally important reason is physical in nature. It was speculated at various occasions that the vacuum might be unstable under domain formation [16]. In the torus geometry this issue can be addressed most efficiently. The main reason is that cubic domains are space-filling. Demonstrating that the vacuum energy density has a minimum at some value of the volume in the torus geometry, is sufficient to establish that the vacuum is unstable with respect to domain formation. It does not necessarily mean that cubic (or for that matter rectangular) domains have the lowest energies, but there is "phenomenological" evidence that the truth might not be far off from that. If we assume the domains to be typically of a size around $(0.8\,\mathrm{fm})^3$, the intermediate volume results imply in a domain-sized vacuum a scalar and tensor glueball mass and a string tension that agree with the Monte Carlo results in volumes beyond $(2\text{–}3\,\mathrm{fm})^3$. For more details we refer to the discussion in refs. [17,18].

Nevertheless, mainly due to the absence of the exact θ-sphaleron and instanton solutions in the torus geometry, which are indispensable guides to the relevant degrees of freedom, attempts to include the θ dependence in this geometry have up to now remained fruitless. We have therefore decided to temporarily shift our interest to a spherical geometry. After all, ultimately the spatial geometry should be irrelevant for the infinite-volume limit. Though we do not expect to be able to consider the infinite-volume limit within our calculational framework, the fact that on S^3 we know the sphaleron and all instanton configurations, will make it a useful laboratory to consider the θ-dependence beyond the semiclassical approximation. Gauge theory in a spherical geometry was studied extensively in the past by Cutkosky and collaborators [7]. At the perturbative level our results agree, the only difference is our way of parametrizing the fields to allow for a more effective description of the sphaleron and instantons. However, non-perturbatively we follow quite a different route.

The remainder of this paper will discuss the technical details of our strategy to include the θ-dependence for the spherical geometry. To keep things transparent we consider pure SU(2) gauge theory. Generalizations to arbitrary gauge groups is in principle (but not necessarily in practice) straightforward. Including chiral fermions (in the fundamental representation of the gauge group) is less straightforward but might present a manageable challenge in the hamiltonian approach when one takes into account that the relevant Dirac vacuum does not respect the symmetry under large gauge transformations [19], as these transformations due to the chiral anomaly [1] do not preserve particle number (θ can indeed be rotated away by a chiral transformation [1,2]). We leave that to future investigations (for a few additional remarks we refer to ref. [20]).

In sect. 2 we give the general SU(2) instanton configurations for the four-dimensional manifold $S^3 \times \mathbb{R}$, obtained by a conformal transformation from the well known instanton solutions on $\mathbb{R}^4$ (or S^4) [21]. We will isolate the sphaleron from this and demonstrate it is indeed only unstable in the tunnelling direction. In sect. 3 we first discuss a toy model to demonstrate that accurate results can be obtained by deriving an effective hamiltonian in the tunnelling degrees of freedom. In its lowest approximation this would for the gauge case lead to a lagrangian in terms of the 5 instanton moduli parameters, much like the description of monopole–monopole scattering [22]. In that case the potential vanishes and the kinetic term follows from the metric of the relevant monopole parameter space, which is known explicitly [23]. Higher-order corrections are then obtained by integrating out all other modes.

In sect. 4 we discuss the hamiltonian version of the streamline or valley equation [24], which needs to be satisfied to integrate out the irrelevant modes. The instanton parameters will not in general satisfy the valley equation. We show that a solution of the valley equation which contains the instanton parameters in lowest order, does not yield a satisfactory result either due to singularities. Among other things, complications arise due to the spherical symmetry of the sphaleron mode. But the single mode that contains this sphaleron does satisfy the valley equation and the perpendicular degrees of freedom can be integrated out consistently. In sect. 5 we show how this is done, avoiding the problem associated with zero eigenvalues for the ghost. From the toy model of sect. 3 one knows that the adiabatic one-loop approximation is governed by a parameter that remains finite as the coupling constant goes to zero. The main reason is that there is no separation of time scales between the tunnelling and perturbative modes in the perturbative regions. All modes should in principle be treated more or less at the same footing. Nevertheless, close to the sphaleron the modes in the tunnelling direction should dominate.

In sect. 6 this hybrid between the perturbative and tunnelling modes is achieved by first considering those 18 perturbative modes that are degenerate in lowest order in perturbation theory and contain all relevant sphalerons and classical vacua. The θ-dependence is included by imposing boundary conditions, similar to what was done for the torus geometry with the constant modes [8]. Our analysis will be based on the intersection of these 18 modes with the boundary of the fundamental modular domain, as was discussed in general terms in ref. [20]. For the torus geometry the constant modes were responsible for an infinite degeneracy in lowest order in perturbation theory. The low-energy dynamics could then be described by an effective hamiltonian [6] in the constant modes, through Bloch's method [25]. The above mentioned boundary conditions could be implemented in a finite-dimensional setting. In the spherical geometry, the modes that carry the boundary conditions do not even have the lowest energy. Here the boundary conditions need to be implemented at the first step in perturbation theory. Rather than taking for the wave functional a product of gaussians, one takes for the 18 modes, that are to lowest order degenerate and carry the sphaleron degrees of freedom, wave functions that suitably incorporate θ-dependence through boundary conditions. All other modes are kept gaussian and one performs perturbation theory as usual. Details of this proposal will be worked out in future publications. In the last section we only give a rough outline and mention some of the issues that require special care. We end with some concluding remarks. Some technical results are collected in appendices A–D.

2. The instantons and sphaleron on $S^3 \times \mathbb{R}$

It is not too difficult to construct all instantons of unit topological charge on $S^3 \times \mathbb{R}$. This is because the self-duality equations are conformally invariant and $S^3 \times \mathbb{R}$ is conformally equivalent to $S^4 = \mathbb{R}^4$. Let $x_\mu \in \mathbb{R}^4$ have the radial decomposition $r^2 = x_\mu^2$, $n_\mu = x_\mu/r$. The conformal equivalence is then specified by redefining time through $r = R\exp(t/R)$ such that

$$dx_\mu^2 = \exp(2t/R)(dt^2 + R^2 dn_\mu^2)\,, \tag{1}$$

where dn_μ^2 represents the metric of the unit three-sphere of volume $2\pi^2$. The vector potential for the instanton is then simply obtained by identifying the connection one-forms

$$A_\mu(t,n)dn_\mu + A_0(t,n)dt = A_\mu dx_\mu|_{x_\mu = R\exp(t/R)n_\mu}\,. \tag{2}$$

We will achieve much simplification in our subsequent computations by using quaternions [26]

$$x = x_\mu \sigma^\mu = \overline{x}^\dagger, \quad \sigma_i = -\overline{\sigma}_i = i\tau_i, \quad \sigma_4 = \overline{\sigma}_4 = 1. \tag{3}$$

The σ_μ are unit quaternions in their 2×2 matrix representation, with τ_i the usual Pauli matrices. Equally useful will be the (anti)self-dual 't Hooft symbols η and $\overline{\eta}$ [1] defined through

$$\begin{aligned} \sigma_\mu\overline{\sigma}_\nu - \sigma_\nu\overline{\sigma}_\mu &= 2i\eta^a_{\mu\nu}\tau_a, \\ \overline{\sigma}_\mu\sigma_\nu - \overline{\sigma}_\nu\sigma_\mu &= 2i\overline{\eta}^a_{\mu\nu}\tau_a. \end{aligned} \tag{4}$$

Unless specified differently, indices $\mu,\nu,\ldots$ will run from 1 to 4 and $a,b,\ldots$ and $i,j,\cdots$ will run from 1 to 3. We also introduce a dreibein on S^3 in terms of η (from now onwards we take $R=1$; R-dependence can be easily reinstated on dimensional grounds)

$$e^a_\mu = \eta^a_{\mu\nu} n_\nu\,. \tag{5}$$

Treating η^a as a four-dimensional antisymmetric orthogonal matrix, it is easily seen from eq. (4) that

$$\eta^a\eta^b = -\varepsilon_{abc}\eta^c - \delta_{ab}\,, \tag{6a}$$

which implies the completeness relations

$$e^a_\mu e^a_\nu + n_\mu n_\nu = \delta_{\mu\nu}, \quad e^a_\mu e^b_\mu = \delta_{ab}\,, \tag{6b}$$

whereas the spin connection ω follows from

$$e^a_\nu \partial_\nu e^b_\mu = \partial_a e^b_\mu = -n_\mu\delta_{ab} + \varepsilon_{abc}e^c_\mu \equiv -n_\mu\delta_{ab} - \omega_{abc}e^c_\mu\,. \tag{7}$$

Flat indices will henceforth be indicated by $i,j,k,\ldots$ to not confuse them with the SU(2) algebra indices.

On $\mathbb{R}^4$ the most general instanton is given by [1,21,26]

$$A_\mu dx_\mu = \mathrm{Im}\left(\frac{(\lambda x + b)\lambda d\overline{x}}{1 + |\lambda x + b|^2}\right), \tag{8}$$

where b_μ and λ are the instanton position and scale parameters. They form the so-called five-dimensional moduli space. Note that our conventions are such that A_μ is antihermitian. A basis of the Lie algebra is provided by $i\tau_a/2 = \frac{1}{2}\sigma_a$. Using eq. (2) one easily obtains the instanton solutions on $S^3 \times \mathbb{R}$ in explicit form:

$$A_0 = \frac{isb\cdot\tau}{1+s^2+b^2+2sb\cdot n}, \quad A_j = -i\frac{(s^2+sb\cdot n)\tau_j + s(b\wedge\tau)_j}{1+b^2+s^2+2sb\cdot n}, \tag{9}$$

where $\boldsymbol{b}$ is the three vector obtained by contracting b_μ with the dreibein e^i_μ, i.e. $\boldsymbol{b} = b\cdot \boldsymbol{e}$ and $s = \lambda e^t$. We observe that the scale parameter of the $\mathbb{R}^4$ (or S^4) instantons is related to the time parameter of the $S^3 \times \mathbb{R}$ instantons. On the other hand, the length of b_μ can be seen as the parameter that describes the size of the instantons on $S^3 \times \mathbb{R}$. At $b = 0$, $A_j = -is^2(1+s^2)^{-1}\tau_j$ and $A_0 = 0$; it represents for each time a constant, rotationally invariant (after a compensating gauge transformation) field configuration. It is convenient to rewrite

$$A_0 = \frac{i\varepsilon\cdot\tau}{2(1+\varepsilon\cdot n)}, \quad \boldsymbol{A} = -i\frac{(u \pm \varepsilon\cdot n)\boldsymbol{\tau} + \boldsymbol{\varepsilon}\wedge\boldsymbol{\tau}}{2(1+\varepsilon\cdot n)}, \tag{10}$$

with

$$u = \frac{2s^2}{1+b^2+s^2}, \quad \varepsilon_\mu = \frac{2sb_\mu}{1+b^2+s^2}. \tag{11}$$

It describes tunnelling from $\boldsymbol{A} = \boldsymbol{0}$ at $t = -\infty$ to $\boldsymbol{A} = -i\boldsymbol{\tau}$ at $t = \infty$. At $t = \infty$ it is a gauge copy of $\boldsymbol{A} = \boldsymbol{0}$ with a gauge transformation $[\Omega]A = \Omega A\Omega^{-1} + \Omega\partial\Omega^{-1}$, where $\Omega = n\cdot\sigma$ is a gauge function with unit winding number. In general this gauge transformation maps instantons to anti-instantons, with u replaced by $2-u$, as follows from

$$n\cdot\sigma\partial_a n\cdot\bar\sigma = -i\tau_a, \quad n\cdot\sigma e^a_\mu\tau_a n\cdot\bar\sigma = \bar e^a_\mu\tau_a, \quad \bar e^a_\mu \equiv \bar\eta^a_{\mu\nu}n_\nu. \tag{12}$$

This is the way time reversal symmetry is implemented on the parameter space. The lagrangian, when restricted to the instantons, must therefore be invariant under $u \to 2-u$. Indeed,

$$\begin{aligned} F_{ij} &= \partial_i A_j - \partial_j A_i + 2\omega_{ijk}A_k + [A_i, A_j] \\ &= \frac{2i\varepsilon_{ijk}\tau_k s^2}{(1+s^2+b^2+2sb\cdot n)^2} \\ &= \frac{i\varepsilon_{ijk}\tau_k(u(2-u)-\varepsilon^2)}{2(1+\varepsilon\cdot n)^2} \end{aligned} \tag{13}$$

yields a potential with the desired symmetry*

$$\mathcal{V} = -\tfrac{1}{2}\int_{S^3}\mathrm{Tr}(F_{ij}^2) = \frac{48\pi^2(1+s^2+b^2)s^4}{((1+b^2+s^2)^2-4s^2b^2)^{\frac{5}{2}}} = \frac{3\pi^2(u(2-u)-\varepsilon^2)^2}{(1-\varepsilon^2)^{5/2}}. \tag{14}$$

* The angular integrations can be reduced to $\int_0^\pi d\psi \sin^2\psi/(p+q\cos\psi)^4$ and can be expressed in terms of $\int_0^\pi d\psi\sin^2\psi/(p+q\cos\psi) = \pi/(p+\sqrt{p^2-q^2})$, which will also be useful further on.

Note that we have not only put $R = 1$, but also $g = 1$, where g is the coupling constant. To restore the proper dependence on the coupling constant the lagrangian is divided by g^2. One also easily verifies the anti-self-duality

$$F_{0i} = \partial_t A_i - \partial_i A_0 + [A_0, A_i] = s\partial_s A_i - \partial_i A_0 + [A_0, A_i] = -\tfrac{1}{2}\varepsilon_{ijk}F_{jk}\,. \tag{15}$$

The change from self-dual to anti-self-dual is because we choose to label x_μ by (x_1, x_2, x_3, x_4), whereas for $S^3 \times \mathbb{R}$ we label time by an index 0.

We now want to identify the sphaleron. In most rigour it is defined in terms of a mini-max procedure. Take any path γ connecting $A_i = 0$ and $A_i = n \cdot \sigma \partial_i n \cdot \overline{\sigma}$ and determine the maximum $\mathcal{V}_{\rm m}(\gamma)$ of the classical potential energy along the path. Taking the minimum of $\mathcal{V}_{\rm m}(\gamma)$ with respect to all γ defines the sphaleron. By construction the sphaleron is a saddle point with precisely one unstable mode. Although not required, it is natural to assume the sphaleron can be found by restricting γ to the instantons. In an obvious notation $\gamma_\varepsilon(u)$ describes a subset of paths, which from eq. (14) have maximal energy $\mathcal{V}_{\rm m}(\gamma_\varepsilon) = 3\pi^2/\sqrt{(1-\varepsilon^2)}$ at $u = 1$. From eq. (11) we see that ε^2 ranges between 0 and 1. The sphaleron is hence expected to coincide with $A_k = -i\tau_k/2$. Its curvature, $F_{ij} = i\varepsilon_{ijk}\tau_k/2$, follows from eq. (13). One easily verifies it is indeed a saddle point for $\mathcal{V}$ on the space of all connections:

$$D_i F_{ij} = \partial_i F_{ij} + \omega_{ijk} F_{ik} + [A_i, F_{ij}] = 0\,. \tag{16}$$

Next we should establish that expansion of $\mathcal{V}$ around this candidate sphaleron has as only unstable mode $\delta A_i = \sigma_i$, corresponding to the tunnelling direction. For later purposes we will slightly generalize this analysis by expanding around $A_k^{(0)} \equiv -iu\tau_k/2$ and $A_0^{(0)} \equiv 0$, which corresponds to $\varepsilon = 0$. We use the background gauge fixing on the variables δA_i,

$$D_i^{(0)}\delta A_i \equiv (\partial_i + {\rm ad}A_i^{(0)})\delta A_i \equiv \partial_i \delta A_i + [A_i^{(0)}, \delta A_i] = 0\,. \tag{17}$$

The fields are considered time-independent and the expansion of $\mathcal{V}$ is identical to that of the full lagrangian, provided $\delta A_0 \equiv 0$ and $\partial_0 \delta A_i \equiv 0$. From eq. (1) we also see that $\mathbb{R}^4$ (remember that we have put $R = 1$) has a vierbein $\tilde{e}^0_\mu = rn_\mu, \tilde{e}^i_\mu = re^i_\mu$, which allows us to immediately copy the stability analysis from 't Hooft's $\mathbb{R}^4$ analysis in the radial representation. We shift t such that $s = r$ and find

$$-\tfrac{1}{2}\int_{S^3} {\rm Tr}\{F_{ij}^2 + 2(D_i^{(0)}\delta A_i)^2\} = \tfrac{1}{2}\int_{S^3} \delta A_i^a \mathcal{M}_{ij}^{ab} \delta A_j^b + {\rm O}(\delta A^3)\,,$$
$$\mathcal{M}_{ij}^{ab} = r^2 e^i_\mu \mathcal{M}_{\mu\nu}^{ab} e^j_\nu\,, \tag{18}$$

where the operator $\mathcal{M}_{\mu\nu}$ is identical to 't Hooft's expression (eq. (2.22) of the first paper in ref. [1]), hence ($u = 2s^2/(1+s^2) = 2r^2/(1+r^2)$)

$$\mathcal{M}_{ij} = \left((2\boldsymbol{L}_1 + u\boldsymbol{T})^2 + 1\right)\delta_{ij} + 2u(2-u)\boldsymbol{T}\cdot\boldsymbol{S}_{ij}\,,$$
$$S^a_{ij} = 2e^i_\mu (S^a_1)_{\mu\nu} e^j_\nu\,, \tag{19}$$

with [1]

$$L_1^a = -\frac{i}{2}\eta^a_{\mu\nu}x^\mu\partial_\nu = \frac{i}{2}\partial_a\,,$$
$$T^a = \mathrm{ad}(\tau_a/2)\,, \qquad T^a_{bc} = -i\varepsilon_{abc}\,,$$
$$(S_1^a)_{\mu\nu} = -\frac{i}{2}\eta^a_{\mu\nu}\,, \qquad S^a_{ij} = -i\varepsilon_{aij}\,. \tag{20}$$

It is also useful to consider [1]

$$L_2^a = -\frac{i}{2}\bar{\eta}^a_{\mu\nu}x^\mu\partial_\nu, \quad (S_2^a)_{\mu\nu} = -\frac{i}{2}\bar{\eta}^a_{\mu\nu}\,, \tag{21}$$

since

$$\boldsymbol{L}_1^2 = \boldsymbol{L}_2^2, \quad 2\bar{e}^i_\mu(S_2^a)_{\mu\nu}\bar{e}^j_\nu = S^a_{ij}, \tag{22}$$

which shows that S^3 still has $SO(4) \simeq (SU(2) \times SU(2))/\mathbb{Z}_2$ symmetry generated by $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, although the SO(4) symmetry of the tangent frames is broken to SO(3). Hence, whereas $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ have spin $\frac{1}{2}$, $\boldsymbol{S}$ has spin 1.

To verify that $\mathcal{M}$ has only one unstable mode, we rewrite it in two different ways:

$$\begin{aligned}\mathcal{M} &= 2(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 + 2(u-1)(\boldsymbol{L}_1 + \boldsymbol{T})^2 + 2(2-u)\boldsymbol{L}_1^2 \\ &\quad + 2(u-1)(2-u)\boldsymbol{T}\cdot\boldsymbol{S} - 2u(2-u)\,,\end{aligned} \tag{23}$$

$$\begin{aligned}\mathcal{M} &= 2u(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 - 2(u-1)(\boldsymbol{L}_1 + \boldsymbol{S})^2 + 2\boldsymbol{L}_1^2 \\ &\quad - 2u(u-1)\boldsymbol{T}\cdot\boldsymbol{S} - 2u(2-u)\,.\end{aligned} \tag{24}$$

The candidate sphaleron corresponds to $u = 1$, where the spectrum of $\mathcal{M}$ is easily derived in explicit form. The singlet with $\boldsymbol{S} + \boldsymbol{T} = \boldsymbol{L}_1 = \boldsymbol{L}_2 = \boldsymbol{0}$, has $\mathcal{M}(u = 1) = -2$ and corresponds to the tunnelling mode. All other modes have $\mathcal{M}(u = 1) > 0$. Note that applying eq. (23) for $u = 0$ and eq. (24) for $u = 2$, one verifies that $\mathcal{M} > 0$ at the classical vacua. Thus $A_i = -i\tau_i/2$ satisfies all criteria for the sphaleron. We can not rigorously exclude that there is no sphaleron with lower energy. However, such a configuration, if it exists, is conjectured to be separated from the classical vacua by an energy barrier much higher than the sphaleron energy of $3\pi^2$ (the dependence on g and R is restored by dividing by g^2R) and need not concern us here. Nevertheless, due to the sphaleron's highly symmetric form it is not unreasonable to assume that there are no sphalerons with less energy.

3. Reduction to the moduli parameters

We first consider a simple toy model to demonstrate that an effective hamiltonian in the tunnelling degrees of freedom can give very accurate results even at energies above the sphaleron energy,

$$H = -\tfrac{1}{2}g^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \tfrac{1}{2}(x^2-1)^2 + 2\omega^2(x)y^2 . \qquad (25)$$

The minima of the potential occur at $x = \pm 1$, $y = 0$ and the sphaleron corresponds to $x = y = 0$, with an energy $\frac{1}{2}$. We can derive an effective hamiltonian in the tunnelling degree of freedom by integrating out the y degree of freedom in an adiabatic approximation

$$H_{\text{eff}} = -\tfrac{1}{2}g^2\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}(x^2-1)^2 + g\omega(x) . \qquad (26)$$

In table 1 we compare our results obtained from a high precision Rayleigh–Ritz analysis for both equations (25) and (26), choosing $\omega(x) = ax^2 + b$. Techniques identical to those described in appendix B of the first paper in ref. [8] were used. Lower bounds were computed as in sect. 5 of the same reference. The energy split of the ground state is compared to the semiclassical prediction as derived from eq. (26), i.e. $\Delta E_0 = 8\sqrt{2g/\pi}\exp(-\frac{4}{3}g^{-1} + 2a)$. We have taken $a = b = 1$ and $a = 3$, $b = 1$. One observes a near perfect agreement, even where the energies are much higher than the sphaleron energy. One might anticipate the semiclassical approximation for the ground-state energy split due to tunnelling, derived from eqs. (25) and (26), to be identical up to relative errors that vanish as some power of g. This is certainly suggested by table 1. However, for $\omega^2(x) = \frac{1}{4} - \nu(\nu+1)(1-x^2)$ it was shown [27] that eq. (25) yields a ground state energy-split of $\Delta E_0 = 8\sqrt{2g(2\nu+1)/\cos(\nu\pi)}\exp(-\frac{4}{3}g^{-1})$, whereas eq. (26) gives $\Delta E_0 = 8[4g^2(1-k)^{1-k}(1+k)^{1+k}]^{1/4}\exp(-\frac{4}{3}g^{-1})$, where $k^2 = 4\nu(\nu+1) > 1$. These results are accurate up to relative errors that vanish as a power of g. Only for k very close to 1 the two results will start to differ significantly. Nevertheless, the relative error in computing the semiclassical result from eq. (26), as compared to eq. (25), does not vanish as a power of g. The reason is that the time scale for the fluctuations in the y-direction is comparable to the time scale for the x-fluctuations and the adiabatic approximation involved in deriving eq. (26) is governed by a parameter that does not vanish as $g \to 0$. To correct for this, higher-derivative terms have to enter the effective hamiltonian. This should in principle be feasible, although it is not very practical.

Still, the most natural strategy would be to derive an effective hamiltonian or lagrangian based on the moduli parameters (s, b_μ) or (u, ε_μ) introduced in the previous section. A priori there is no reason to suspect the adiabatic approximation to be bad, since at $A = 0$ all modes are quadratic (unlike in the torus geometry [6,8]), as can be seen from eq. (24) at $u = 0$. In lowest order the effective lagrangian is obtained by restricting the full Yang–Mills action $S = -\frac{1}{2}\int \mathrm{d}t \int_{S^3} \mathrm{Tr}(F_{\mu\nu}^2)$ to the instanton moduli space, eq. (9) or (10), where the parameters (s, b_μ) and (u, ε_μ) are now considered as arbitrary functions

TABLE 1

The spectrum for the hamiltonian of eq. (25) [denoted by f in the first column], with $\omega(x) = ax^2 + b$, as compared to the spectrum for the effective hamiltonian of eq. (26) [denoted by e]. Energy levels are labelled by the parity under $x \to -x$ and the last column compares the ground-state energy split $\Delta E_0 \equiv E_0^- - E_0^+$ with the semiclassical result derived from eq. (26), i.e. WKB $= 8(2g/\pi)^{1/2} \exp(-\frac{4}{3}g^{-1} + 2a)$. All the digits displayed are significant.

g	E_0^+	E_1^+	E_2^+	E_0^-	E_1^-	E_2^-	$\Delta E_0/\text{WKB}$
			$a = 1$, $b = 1$				
0.07f	0.203913	0.329881	0.444434	0.203913	0.329890	0.444933	0.824
0.07e	0.203674	0.329806	0.444497	0.203674	0.329815	0.444999	0.828523
0.08f	0.232003	0.373217	0.496612	0.232004	0.373295	0.499910	0.8010
0.08e	0.231693	0.373121	0.496681	0.231694	0.373200	0.499995	0.804366
0.10f	0.287343	0.455645	0.581230	0.287361	0.457204	0.608842	0.754446
0.10e	0.286864	0.455504	0.581303	0.286883	0.457068	0.608999	0.756140
0.50f	1.07977	2.06444	3.46392	1.44538	2.79115	4.40555	0.157774
0.50e	1.07362	2.06399	3.62257	1.43613	2.79554	4.52717	0.156437
1.00f	2.05168	5.20259	7.06059	3.43401	7.28025	1.01147	0.111187
1.00e	2.03018	5.22785	9.63091	3.39984	7.32237	1.21192	0.110167
			$a = 3$, $b = 1$				
0.07f	0.319200	0.432434	0.526924	0.319201	0.432632	0.533333	0.4859
0.07e	0.318516	0.431953	0.526706	0.318518	0.432152	0.533135	0.495291
0.08f	0.359561	0.482468	0.571896	0.359579	0.483967	0.595335	0.43398
0.08e	0.358657	0.481854	0.571720	0.358676	0.483360	0.595168	0.440739
0.10f	0.435852	0.564875	0.661462	0.436300	0.581781	0.729033	0.339584
0.10e	0.434417	0.564106	0.661392	0.434868	0.580958	0.729042	0.342075
0.50f	1.32849	2.84047	3.95723	2.04007	3.77051	5.63556	0.562421
0.50e	1.31046	2.84923	4.78421	2.01298	3.77879	5.85457	0.555258
1.00f	2.64386	7.18664	8.55272	4.96345	1.00281	1.37669	0.341722
1.00e	2.57941	7.39351	1.29745	4.86979	1.01043	1.59855	0.337419

of time. Choosing b_μ time independent and $s = \lambda e^t$ will indeed describe the instanton solution of the truncated lagrangian, as we will demonstrate now.

We have already computed the potential in eq. (14) and we are left with computing the kinetic term. As the moduli space is a gauge invariant object we should eliminate all gauge degrees of freedom. The kinetic term is therefore determined by the metric on the moduli space, which follows from truncating the riemannian metric of the full configuration space. If s is the invariant line element one has [28]

$$\dot{s}^2 = -\int_{S^3} \mathrm{Tr}(\dot{A}_i - D_i D_j^{-2} D_k \dot{A}_k)^2 , \tag{27}$$

where $D_i = \partial_i + \mathrm{ad}A_i$ is the covariant derivative and D_j^{-2} denotes the Green function for the covariant laplacian D_j^2 (our indices are flat indices and that the spin connection will not enter in eq. (27)). Alternately, this can be written as

$$\dot{s}^2 = -\int_{S^3} \mathrm{Tr}(\dot{A}_i^2 + D_i\dot{A}_i D^{-2} D_k \dot{A}_k)\,. \tag{28}$$

For the moduli space

$$D_i\dot{A}_i = \frac{3i(u-1)\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}}{2(1+\varepsilon\cdot n)^2} - \frac{i\dot{u}(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau})}{2(1+\varepsilon\cdot n)^2} - \frac{i(u-1)\dot{\varepsilon}\cdot n\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}}{(1+\varepsilon\cdot n)^3} - \frac{i\boldsymbol{\varepsilon}\cdot(\dot{\boldsymbol{\varepsilon}}\wedge\boldsymbol{\tau})}{2(1+\varepsilon\cdot n)^2}\,. \tag{29}$$

Unfortunately, all our attempts to exactly invert the covariant laplacian have failed. Nevertheless, some explicit information can be extracted by using eq. (15), which implies

$$sD_i(\partial_s A_i) = D_i^2 A_0\,. \tag{30}$$

Writing the riemannian metric in terms of the moduli parameters s and b_μ this gives

$$\begin{aligned} \dot{s}^2 &\equiv g_{00}\dot{s}^2 + 2g_{0\mu}\dot{s}\dot{b}_\mu + g_{\mu\nu}\dot{b}_\mu\dot{b}_\nu\,,\\ g_{00} &= -\int_{S^3}\mathrm{Tr}\{(\partial_s A_i)^2 + s^{-1}(D_i\partial_s A_i)A_0\} = \frac{\mathcal{V}}{s^2}\,,\\ g_{0\mu} &= -\int_{S^3}\mathrm{Tr}\{(\partial_s A_i)(\partial_{b_\mu}A_i) + s^{-1}(D_i\partial_{b_\mu}A_i)A_0\} = \frac{-2\mathcal{V}b_\mu}{s(s^2+b^2+1)}\,, \end{aligned} \tag{31}$$

with $\mathcal{V}$ the potential as given in eq. (14). Evaluation of eq. (31) involves angular integrals similar to those encountered in evaluating eq. (14). (See the remarks made there.) The metric components $g_{\mu\nu}$ can only be computed in an expansion in powers of b, to which we will return in sect. 4. We can, however, with this result verify that $\dot{b}_\mu = 0$ and $s = \lambda e^t$ is indeed an instanton solution for the truncated moduli lagrangian (with time imaginary)

$$L_{\mathrm{mod}} = g_{00}\dot{s}^2 + 2g_{0\mu}\dot{s}\dot{b}_\mu + g_{\mu\nu}\dot{b}_\mu\dot{b}_\nu + \mathcal{V}\,, \tag{32}$$

since we can split off squares:

$$\begin{aligned} L_{\mathrm{mod}} &= \mathcal{V}\left(\frac{\dot{s}}{s} - \frac{2\dot{b}_\mu b_\mu}{(1+b^2+s^2)} \pm 1\right)^2\\ &\quad + \left(g_{\mu\nu} - \frac{4\mathcal{V}b_\mu b_\nu}{(1+b^2+s^2)^2}\right)\dot{b}_\mu\dot{b}_\nu \pm 8\pi^2\dot{Q}\,, \end{aligned} \tag{33}$$

where Q is defined through

$$\frac{\partial Q}{\partial s} = \frac{\mathcal{V}}{4\pi^2 s}\,, \qquad \frac{\partial Q}{\partial b_\mu} = -\frac{b_\mu\mathcal{V}}{2\pi^2(1+b^2+s^2)}\,. \tag{34}$$

Using the explicit form of $\mathcal{V}$ in eq. (14), one easily verifies that Q is integrable. Fixing an irrelevant additive constant through $Q(0,0) = 0$, integration yields

$$Q(s,b) = \tfrac{1}{2} + \frac{s^2 - 1 - b^2}{\sqrt{(1 + b^2 + s^2)^2 - 4s^2b^2}} \left\{ \tfrac{1}{2} + \frac{s^2}{(1 + b^2 + s^2)^2 - 4s^2b^2} \right\}, \tag{35}$$

from which one finds, as it should be, that $Q(\infty,b) - Q(0,b) = 1$, which is the instanton "charge". One also verifies that $Q(1,0) = \frac{1}{2}$. The "charge" of the sphaleron relative to the vacuum is half the instanton "charge" [4]. Finally, it should be needless to point out that $\dot{b}_\mu = 0$, $s = \lambda e^{\pm t}$ *uniquely* solves the equations of motion of L_{mod}.

As mentioned before, we have been unable to compute $g_{\mu\nu}$ exactly. For metrics on moduli spaces quite a lot is known. Their significance in the physics of gauge theories was first recognized in ref. [22] for the monopole scattering problem. In that case (in the BPS limit) there is no potential energy and all the dynamics is determined by the metric. Due to its large symmetry and the use of a hyperKähler structure, the exact metric (a four-dimensional self-dual Einstein manifold not known before) could be constructed in terms of elliptic functions. Earlier, Donaldson had been using instanton moduli space to study the differential geometry of four-manifolds [29]. The relevant metric in the four-dimensional context is of the same form as eq. (27), except that we now integrate over the four-dimensional compact manifold and the indices $i, j, k, \ldots$ run from 1 to 4. Although eq. (30) is also valid in this four-dimensional context of $S^3 \times \mathbb{R}$, allowing one to compute the g_{00} and $g_{0\mu}$ components of the metric, failure to construct the exact Green function for the covariant laplacian prevented us from computing $g_{\mu\nu}$ in the four-dimensional context too. The Green function is required to project on the horizontal (transverse) directions. This projection is not preserved under conformal transformations. For $\mathbb{R}^4$ the semiclassical computation is considerably simplified using the fact that position parameters are related to translations and the scale parameter to scale transformations, avoiding explicit use of a metric on the moduli space. The semiclassical result [1,5], however, suffers from infrared divergences in the integration over the scale parameter. On the other hand, the semiclassical instanton calculation for S^4 is well defined [30]. In that case, an SO(5) symmetry of the moduli space avoids the explicit use of the metric. Later, the same symmetry was also crucial for constructing the explicit form of the metric [31,32], independently used to derive the semiclassical instanton contribution [31] (it would be a useful check to verify if both results agree). It is not excluded, however, that for the $S^3 \times \mathbb{R}$ geometry it is impossible to compute the semiclassical instanton contribution in closed form.

Nevertheless, a lot is implicitly known about the metric properties of moduli space for a large class of four-dimensional manifolds [29]. In particular there do in general occur points with curvature singularities (the CP_2-cones) associated to reducible connections (these are connections left invariant by a non-trivial subgroup of the gauge group). In the four-dimensional context quite detailed information about the nature of the singularity is available [33]. Both for the three-dimensional monopole and four-dimensional instanton bundles, the field strength is non-vanishing, in which case these singularities are relatively mild [29]. As $\boldsymbol{A} = \boldsymbol{0}$ and its gauge copies are reducible connections, we should anticipate curvature singularities at these points ($u = \varepsilon = 0$ and $u = 2, \varepsilon = 0$), which will be investigated in sect. 4.

4. The valley equation

Our strategy so far has been to include all moduli parameters in the effective theory, to stay as close as possible to the parameterization used in the semiclassical computation. As we have seen, this is partly guaranteed by the fact that the instantons are also part of the effective theory in this approach. But it is necessary to be able to integrate out all other modes, different from the moduli parameters, consistently. As our background [eqs. (9) and (10)] is not a solution of the equations of motion, this requires special care. Due to the arbitrary time dependence of the moduli parameters we should demand the quantum modes to be chosen such that both the kinetic and the potential parts contain no terms linear in the quantum modes. Thus, writing $A_i = A_i(u, \varepsilon) + q_i$ and choosing the background gauge (eq. (17)) $D_i q_i = (\partial_i + \mathrm{ad} A_i(u, \varepsilon)) q_i = 0$, $\dot{q}_i$ needs to be perpendicular to $\dot{A}_i$, i.e. $\int_{S^3} \mathrm{Tr}(\dot{A}_i \dot{q}_i) = 0$. As the time dependence is arbitrary this implies q_i has to be perpendicular to the tangents of the moduli space

$$\int_{S^3} \mathrm{Tr}\left(q_i \frac{\partial A_i}{\partial u}\right) = \int_{S^3} \mathrm{Tr}\left(q_i \frac{\partial A_i}{\partial \varepsilon_\mu}\right) = 0. \tag{36}$$

The linear term in the potential part is proportional to $\int_{S^3} \mathrm{Tr}(q_i D_i F_{ij})$, which in the light of eq. (36) and the background gauge condition can only vanish simultaneously if

$$D_i F_{ij} = \lambda_0 \frac{\partial A_i}{\partial u} + \lambda_\mu \frac{\partial A_i}{\partial \varepsilon_\mu} + D_i \Lambda, \tag{37}$$

where λ_0 and λ_μ depend exclusively on the moduli parameters, whereas Λ is an arbitrary Lie algebra valued function on S^3. This equation is well known as the valley or the streamline equation [24], developed in the four-dimensional context in order to consistently expand around configurations that are not stationary. We therefore consider eq. (37) as the hamiltonian version of the valley equation.

Apart from removing the term linear in the quantum modes it is also important that the zero-point energy of the quantum fluctuations is positive for all quantum modes. Equivalently, the effective hamiltonian should contain all modes that would be unstable somewhere. The notion of stability is obscured, however, by the fact that the spectrum of the quadratic part of the energy functional (the hessian for $\mathcal{V}$) depends on the choice of coordinates, unless it belongs to a stationary point. This is why the notion of an unstable manifold [34] is so useful. Unstable manifolds are associated to saddle points and are obtained by following the gradient lines in the unstable directions. As the instantons are exactly devised to follow these gradients, the subset $(u, \varepsilon = 0)$ of the moduli space coincides with the unstable manifold associated with the sphaleron. It seems, however, when expanding the potential $\mathcal{V}$ on the moduli space with respect to ε,

$$\mathcal{V}(u,\varepsilon) = 3\pi^2 u^2(2-u)^2 + \tfrac{3}{2}\pi^2 u(2-u)(1-5(u-1)^2)\varepsilon^2 + \mathrm{O}(\varepsilon^4)\,, \tag{38}$$

that the ε-direction becomes unstable for $(u-1)^2 < \frac{1}{5}$. As we pointed out, this can be misleading, as the spectrum of the hessian is only well defined (i.e coordinate independent) at $u = 0,\ 1,\ 2$. Indeed the configurations $A_i(u,0) = -iu\tau_i/2$ satisfy the valley equation, as $D_iF_{ij} = -iu(u-1)(u-2)\tau_j$ (in eq. (37), take $\lambda_\mu = 0$, $\Lambda = 0$ and $\lambda_0 = 2u(u-1)(u-2)$). One can integrate out all other modes consistently if the quadratic fluctuation operator $\mathcal{M}(u)$ constructed in eq. (19) (see also eqs. (23) and (24)) is positive definite in the subset of modes that are perpendicular to the tunnelling mode, $\delta A_i = \sigma_i$. It is proven in appendix A that this is indeed the case. There we also show that the non-linearity of $A(u,\varepsilon)$ is responsible for the seemingly additional unstable behaviour in the ε direction in eq. (38). In our subsequent analysis this will also become clear without fixing the gauge.

We will now investigate whether we can, nevertheless, consistently take the ε-mode into account in the effective theory. The rationale being to try and stay as close as possible to the semiclassical result, despite the fact that our toy model at the beginning of sect. 3 has shown that a simple one-loop correction will not be sufficient to reproduce the semiclassical energy split up to relative errors that vanish as $g \to 0$. It will, however, provide important information about the embedding of the tunnelling and sphaleron configurations in the configuration space. Thus we need to investigate whether the moduli parameters satisfy the valley equation. Observe that Λ can be easily determined in terms of the parameters λ_0 and λ_μ, since $D_j(D_iF_{ij}) = -\frac{1}{2}[F_{ij}, F_{ij}] = 0$, such that

$$D_j^2\Lambda = -\lambda_0 D_j(\partial_u A_j) - \lambda_\mu D_j(\partial_{\varepsilon_\mu} A_j)\,. \tag{39}$$

Solving for Λ requires inverting the covariant laplacian, which we were unable to do in closed form. Hence we will solve the valley equation by expanding around $A(u,0)$ in powers of ε.

First we note that the valley equation has two symmetries. Namely local field dependent gauge transformations and coordinate transformations on the moduli parameters: $(u,\varepsilon) \to (u'(u,\varepsilon), \varepsilon'(u,\varepsilon))$, with the constraint that $u'(u,0) = u$, $\varepsilon'(u,0) = 0$. Under such a coordinate transformation, Λ transforms like a scalar and λ_α ($\alpha = 0,\ldots,4$) as a vector with respect to the moduli space coordinates. The λ_α, however, remain unchanged under a gauge transformation $\widetilde{A}_i = [\Omega]A_i \equiv \Omega A_i \Omega^{-1} + \Omega\partial_i\Omega^{-1}$, whereas

$$\widetilde{\Lambda} = \Omega\Lambda\Omega^{-1} - \lambda_0\Omega\partial_u\Omega^{-1} - \lambda_\mu\Omega\partial_{\varepsilon_\mu}\Omega^{-1} . \tag{40}$$

The most general expansion around $A(u,0)$ is given by

$$\boldsymbol{A} = if\boldsymbol{\tau} + ig(\boldsymbol{\varepsilon}\wedge\boldsymbol{\tau}) + ih\boldsymbol{\varepsilon}(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}), \tag{41}$$

with (expanding $\boldsymbol{A}$ to third order in ε)

$$\begin{aligned}
f &= -\tfrac{1}{2}u + f_{-1}(\varepsilon\cdot n) + f_0(\varepsilon\cdot n)^2 + \widetilde{f}_0(\varepsilon\cdot\varepsilon) \\
&\quad + f_1(\varepsilon\cdot n)^3 + \widetilde{f}_1(\varepsilon\cdot n)(\varepsilon\cdot\varepsilon) + \mathrm{O}(\varepsilon^4), \\
g &= g_{-1} + g_0(\varepsilon\cdot n) + g_1(\varepsilon\cdot n)^2 + \widetilde{g}_1(\varepsilon\cdot\varepsilon) + \mathrm{O}(\varepsilon^3), \\
h &= h_0 + h_1(\varepsilon\cdot n) + \mathrm{O}(\varepsilon^2), \\
\lambda_0 &= \lambda_{(1)} + (\varepsilon\cdot\varepsilon)\lambda_{(2)} + \mathrm{O}(\varepsilon^4), \\
\lambda_\mu &= (\lambda_{(3)} + (\varepsilon\cdot\varepsilon)\lambda_{(4)})\varepsilon_\mu + \mathrm{O}(\varepsilon^5), \\
\Lambda &= \Lambda_1 + (\varepsilon\cdot n)\Lambda_2 + (\varepsilon\cdot n)^2\Lambda_3 + (\varepsilon\cdot\varepsilon)\Lambda_4 + \mathrm{O}(\varepsilon^3).
\end{aligned} \tag{42}$$

To third order there are 19 unknown coefficients, that are functions of u only. Since D_iF_{ij}, in terms of $\widehat{f}$, $\widehat{g}$ and $\widehat{h}$ (being rather complicated functions of f, g and h), has the same form as eq. (41),

$$D_iF_{ij} = i\widehat{f}\tau_j + i\widehat{g}(\boldsymbol{\varepsilon}\wedge\boldsymbol{\tau})_j + i\widehat{h}\varepsilon_j(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}), \tag{43}$$

we have one more equation than there are parameters in f, g and h. The additional equation arises since $f(\varepsilon = 0) = -\frac{1}{2}u$ was fixed. This leaves seven coefficients undetermined, which are however removed by the seven independent parameters that describe the gauge and coordinate invariance of our ansatz to the desired order in ε:

$$\begin{aligned}
u' &= u + (\varepsilon\cdot\varepsilon)\theta_1 + \mathrm{O}(\varepsilon^4), \\
\varepsilon'_\mu &= (\theta_2 + (\varepsilon\cdot\varepsilon)\theta_3)\varepsilon_\mu + \mathrm{O}(\varepsilon^5), \\
\Omega &= \exp(i\Theta\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}), \\
\Theta &= \theta_4 + (\varepsilon\cdot n)\theta_5 + (\varepsilon\cdot n)^2\theta_6 + (\varepsilon\cdot\varepsilon)\theta_7 + \mathrm{O}(\varepsilon^3).
\end{aligned} \tag{44}$$

It is left to the reader to verify that taking

$$\theta_2 = \frac{-2(g_{-1} + (u-1)f_{-1})}{u(2-u)}, \qquad \theta_4 = \frac{-(f_{-1} + (u-1)g_{-1})}{u(2-u)} \tag{45}$$

implies that eq. (41) agrees with $A(u, \varepsilon)$ up to first order in ε, or

$$f_{-1} = \tfrac{1}{2}(u-1), \quad g_{-1} = -\tfrac{1}{2}. \tag{46}$$

Note that this might introduce coordinate singularities at $u = 0$ or 2. For the choice of eq. (46) the transformation is appropriately inert. In case $g_{-1} + (u-1)f_{-1} = 0$, eq. (45) should not be used. However, since in the latter case eq. (41) corresponds to $[\Omega]A(u,0)$ to linear order in ε (with $\Omega = \exp[if_{-1}(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau})]$), ε_μ would *not* represent an additional degree of freedom, such that we can safely ignore this case. One can eliminate the five additional redundant parameters as follows: First choose θ_1 to eliminate $\tilde{f}_0$, then θ_5 to remove h_0, subsequently we pick θ_6, θ_7 and θ_3 to eliminate h_1, $\tilde{g}_1$ and $\tilde{f}_1$.

After considerable but straightforward (computer-assisted) algebra we can solve all $\lambda_{(i)}$ and A_i algebraically in terms of the remaining f_i and g_i. For the latter, there remain four equations linear in f_i and g_i, and their first derivatives, which can all be solved explicitly, albeit in closed form only for f_0 and g_0. In appendix B we will give these explicit solutions. Here we will be satisfied with discussing the result for f_0 and g_0, as our subsequent arguments, partly due to some good fortune, will not require the explicit solutions for f_1 and g_1. Introducing $a = u - 1$ (such that time reversal corresponds to $a \to -a$) f_0 and g_0 are determined by the equations

$$\begin{aligned} 2a(1-a^2)\frac{\mathrm{d}f_0(a)}{\mathrm{d}a} + 2(2-a^2)f_0(a) + 8ag_0(a) &= 3a^3, \\ 2a(1-a^2)\frac{\mathrm{d}g_0(a)}{\mathrm{d}a} + 2(2-3a^2)g_0(a) + 4af_0(a) &= \tfrac{3}{2}(1-3a^2). \end{aligned} \tag{47}$$

Comparing eqs. (41), (42) and (10), the moduli space configuration corresponds to (note that it has indeed $f_{-1} = \frac{1}{2}(u-1)$, $g_{-1} = -\frac{1}{2}$, $\tilde{f}_0 = \tilde{f}_1 = \tilde{g}_1 = h_i = 0$)

$$f_0(a) = -\tfrac{1}{2}a, \quad g_0(a) = \tfrac{1}{2} \tag{48}$$

and one immediately verifies that it does not satisfy eq. (47). The important conclusion is therefore that *the moduli space connection does not satisfy the valley equation.*

Nevertheless, a solution of the valley equation can be found (it can be shown that this can be extended to arbitrary order in ε). Explicitly, eq. (47)

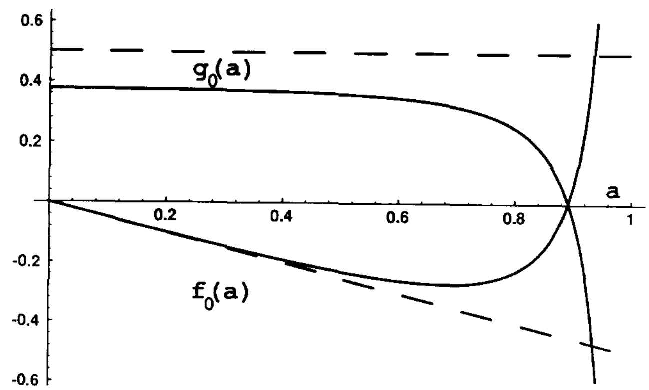

Fig. 1. The functions $f_0(a)$ and $g_0(a)$ for the solution of the valley equation, given in eq. (49) (full curves), compared to their values for the moduli space, eq. (48) (dashed curves). At the scale of the figure it is not visible that both curves do *not* change sign for the same value of a.

is solved by

$$g_0(a) = \frac{c_1(1+3a^2)}{a^2(1-a^2)^{3/2}} - \frac{(47+70a^2+3a^4)}{12a^2(1-a^2)} - \frac{(3+a^2)(5\arcsin a - c_2)}{2a(1-a^2)^{3/2}},$$

$$f_0(a) = \frac{(143-32a^2+9a^4)}{12a(1-a^2)} - \frac{4c_1}{a(1-a^2)^{3/2}} + \frac{(3+6a^2-a^4)(5\arcsin a - c_2)}{4a^2(1-a^2)^{3/2}}, \tag{49}$$

where c_1 and c_2 are constants of integration. It coincides with eq. (48) at $a^2 = 1$ if we choose $c_1 = 5\pi/2$ and $c_2 = 0$, in which case both $f_0(a)$ and $g_0(a)$ become singular at $a = 0$. To make $f_0(a)$ and $g_0(a)$ regular at $a = 0$, one is required to choose $c_1 = \frac{47}{12}$ and $c_2 = 0$, in which case $f_0(a) = -\frac{1}{2}a + O(a^3)$ and $g_0(a) = \frac{3}{8} + O(a^2)$, with singular behaviour at $a^2 = 1$. For the latter case, f_0 and g_0 are plotted in fig. 1, together with the moduli space expressions of eq. (48). It will depend on the nature of these singularities if the effective theory based on the solution of the valley equation presents a well defined reduction. In particular, these singularities might be merely coordinate singularities, possibly introduced through eq. (45). As we have discussed in sect. 3, we do indeed anticipate singularities at $a^2 = 1$ ($u = 0$ and 2). The only safe way to decide on the nature of the singularities is to compute the riemannian curvature tensor. Since R^2_{ijkl} is the sum of squares (contraction over all *flat* indices i, j, k, l is implied), each R_{ijkl} is finite iff R^2_{ijkl} is finite. This also implies that any other scalar invariant will remain finite. We will compute the curvature for arbitrary u at $\varepsilon = 0$, for which it is sufficient to know g_{00}, $g_{0\mu}$ and $g_{\mu\nu}$ to second order in ε. Appendix C will provide some of the details involved in the computation, especially in expanding the inverse covariant laplacian. Here we only list the results in terms of the functions f_i and g_i, such that they can be used for both the moduli and the valley case. In the following we prefer to use (a, ε_μ) as coordinates:

$$\begin{aligned}
\dot{s}^2 &= g_{00}\dot{a}^2 + 2g_{0\mu}\dot{a}\dot{\varepsilon}_\mu + g_{\mu\nu}\dot{\varepsilon}_\mu\dot{\varepsilon}_\nu\,,\\
g_{00} &= 3\pi^2\left\{1 + (\varepsilon\cdot\varepsilon)\left(\frac{a^2}{2(1+2a^2)} - \frac{\mathrm{d}f_0(a)}{\mathrm{d}a}\right)\right\} + \mathrm{O}(\varepsilon^4)\,,\\
g_{0\mu} &= 3\pi^2\left\{\frac{a(2+a^2)}{2(1+2a^2)} - f_0(a)\right\}\varepsilon_\mu + \mathrm{O}(\varepsilon^3)\,,\\
g_{\mu\nu} &= 2\pi^2\left\{\delta_{\mu\nu}\left(\frac{3(1-a^2)^2}{4(1+2a^2)} + (\varepsilon\cdot\varepsilon)\mathcal{G}_1\right) + \varepsilon_\mu\varepsilon_\nu\mathcal{G}_2\right\} + \mathrm{O}(\varepsilon^4)\,,
\end{aligned} \tag{50}$$

where

$$\begin{aligned}
\mathcal{G}_1 &= \frac{\left[\tfrac{1}{8}(-3 - a^2 + 89a^4 - 13a^6) + \tfrac{1}{2}a(29 - 17a^2)(1+a^2)f_0(a) + a^2(a^2-7)(1+3a^2)g_0(a)\right]}{(3+a^2)(1+2a^2)^2}\\
&\quad + \frac{\tfrac{1}{3}(5+3a^2)f_0^2(a) + \tfrac{8}{3}af_0(a)g_0(a) + 4g_0^2(a)}{(3+a^2)}\\
&\quad - \frac{(1-a^2)(3f_1(a) + g_1(a))}{2(1+2a^2)}\,,\\
\mathcal{G}_2 &= \frac{\left[\tfrac{1}{8}(3 + 100a^2 + 55a^4 - 14a^6) - a(1+a^2)(13+11a^2)f_0(a) + a^2(-17 - a^2 + 6a^4)g_0(a)\right]}{(3+a^2)(1+2a^2)^2}\\
&\quad + \frac{\tfrac{2}{3}(8+3a^2)f_0^2(a) + \tfrac{4}{3}af_0(a)g_0(a) + 2g_0^2(a)}{(3+a^2)}\\
&\quad - \frac{(1-a^2)(3f_1(a) + g_1(a))}{(1+2a^2)}\,.
\end{aligned} \tag{51}$$

Before presenting the riemannian curvature there are two important observations. First, the transformation

$$\varepsilon'_\mu = \left\{1 + \frac{3f_1(a) + g_1(a)}{3(1-a^2)}(\varepsilon\cdot\varepsilon)\right\}\varepsilon_\mu \tag{52}$$

allows us to remove f_1 and g_1 from the metric to the relevant order in ε; they will thus not appear in the curvature. Secondly, the transformation

$$a = u - 1 = a' + \tfrac{1}{2}\left\{ f_0(a') - \frac{a'(2 + a'^2)}{2(1 + 2a'^2)} \right\} (\varepsilon \cdot \varepsilon) \tag{53}$$

removes to this order the term $g_{0\mu}$. One easily verifies that for general f_i and g_i

$$\mathcal{V}(a, \varepsilon) = 3\pi^2(1 - a^2)^2 + \frac{3\pi^2}{2}(1 - a^2)(1 - 3a^2 + 4af_0(a))(\varepsilon \cdot \varepsilon) + \mathrm{O}(\varepsilon^4), \tag{54}$$

which for $u = a + 1$, $f_0(a) = -\frac{1}{2}a$ coincides with the moduli space result of eq. (38). Remarkably, eq. (53) is seen to transform the potential into

$$\mathcal{V}(a', \varepsilon) = 3\pi^2(1 - a'^2)^2 + \frac{3\pi^2}{2}(1 - a'^2)^2 \frac{1 + 4a'^2}{1 + 2a'^2}(\varepsilon \cdot \varepsilon) + \mathrm{O}(\varepsilon^4), \tag{55}$$

which is independent of f_i and g_i to the displayed order and has a positive coefficient for the part quadratic in ε. We thus confirm that the ε mode can be integrated out quadratically, in agreement with the analysis in appendix A.

Finally we present the curvature at $\varepsilon = 0$. It has only two independent components,

$$R_{m0n0} = \delta_{mn} R_I, \qquad R_{mnpq} = (\delta_{mq}\delta_{np} - \delta_{mp}\delta_{nq}) R_{II}, \tag{56}$$

all others vanish if not related to these by the usual symmetries in permuting the indices. The indices m, n, p, q run from 1 to 4; together with 0 they form the *flat* indices associated to the metric of eq. (50). In our conventions a sphere has a positive Ricci scalar $R \equiv R_{pqpq}$. The quantities R_I and R_{II} are given by

$$\begin{aligned} R_I &= \frac{4}{(1 + 2a^2)^2}, \\ R_{II} &= \frac{2(-3 + 62a^2 + 34a^4 + 87a^6 + 157a^8 - 16a^{10})}{3(1 - a^2)^4(3 + a^2)(1 + a^2)^2} \\ &\quad + \frac{8a(2 - a^2)(5 + 7a^2)f_0(a) + 16a^2(1 - 13a^2)g_0(a)}{3(1 - a^2)^4(3 + a^2)} \\ &\quad + \frac{(8(5 + 3a^2)f_0(a)^2 + 64af_0(a)g_0(a) + 96g_0(a)^2)(1 + 2a^2)^2}{9(1 - a^2)^4(3 + a^2)}. \end{aligned} \tag{57}$$

For the moduli space, eq. (48) implies

$$R_{II} = \frac{2(1 + 79a^2 + 149a^4 + 88a^6 + 16a^8)}{3(1 - a^2)^2(3 + a^2)(1 + 2a^2)^2}. \tag{58}$$

The conclusion is that both the moduli and valley case have real curvature singularities at $a^2 = 1$, i.e. at the perturbative vacua. The valley case will furthermore have a curvature singularity at $a = 0$, unless we take in eq. (49) $c_1 = \frac{47}{12}$ and $c_2 = 0$. But for that choice, the curvature singularity at $a^2 = 1$ is proportional to $(1 - a^2)^{-7}$.

A singularity at $A = 0$ was anticipated, due to the fact that the gauge group has been divided out. For example in the Coulomb gauge $\partial_i A_i = 0$, the hamiltonian [35] is regular at $A = 0$, because the constant gauge transformations are not fixed. Instead, one demands the wave functional to be a singlet under this remaining gauge symmetry. We could prove that neither in the moduli, nor in the valley case, is it possible to remove the singularity in this way. One source of additional singular behaviour arises due to the rotational invariance of the tunnelling path with $\varepsilon = 0$. This will be demonstrated in sect. 5 by embedding it in the constant modes (i.e. $\boldsymbol{L}_1^2 = 0$). Another source is, that at $a^2 = 1$, the ε mode is a pure gauge (explaining why $\mathcal{V}$ is quartic, rather than quadratic in ε, at $a^2 = 1$). We should thus conclude that, mainly due to non-linearities and a singular embedding of the tunnelling path, including the ε mode is an ill-posed problem. However, it is consistent to expand around the one-parameter tunnelling path through the sphaleron ($\varepsilon = 0$), henceforth to be called the *sphaleron path*, as it satisfies the hamiltonian valley equation and all zero-point frequencies for the fluctuations perpendicular to this sphaleron path are real. In the next section we will show how one deals in this framework with the coordinate singularity at the classical vacua.

5. Reduction to the sphaleron mode

In this section we analyse the embedding of the unstable manifold, $A_i = -iu/2\tau_i$, within the set of modes with $\boldsymbol{L}_1^2 = 0$, which we will call *constant modes*. They are labelled by

$$A_i = ic_i^a \tau_a / 2. \tag{59}$$

This subsector was analysed before [12,36] by decomposing c_i^a in three gauge modes, three angular modes and three "radial" modes. We follow the notation of Koller and van Baal [12],

$$c_i^a = \sum_{b=1}^{3} \xi^{ab} x_b \eta_{bi}, \tag{60}$$

where $\xi, \eta \in SO(3)$. If $\widehat{T}_i$ are the gauge generators associated to ξ and $\widehat{L}_i$ the rotation generators belonging to η, one finds the hamiltonian restricted to the constant modes to be

$$\mathcal{H}(c) = -\frac{g^2}{4\pi^2}\frac{\partial^2}{\partial c_i^{a2}} + \frac{\mathcal{V}(c)}{g^2}$$
$$= -\frac{g^2}{4\pi^2}\left\{\frac{1}{\mathcal{J}}\frac{\partial}{\partial x_i}\mathcal{J}\frac{\partial}{\partial x_i} - \frac{1}{4}\sum_{i\neq j\neq k}\left(\frac{\widehat{L}_k + \widehat{T}_k}{x_i - x_j}\right)^2 + \left(\frac{\widehat{L}_k - \widehat{T}_k}{x_i + x_j}\right)^2\right\}$$
$$+\frac{\mathcal{V}(x)}{g^2}, \tag{61}$$

with $\mathcal{J}$ the jacobian associated to the change of variables. Explicitly

$$\mathcal{J} = \prod_{i>j} |x_i^2 - x_j^2|. \tag{62}$$

To avoid double counting, one restricts x_i to the double cone (or diabolo)

$$0 \leqslant \pm x_1 \leqslant \pm x_2 \leqslant \pm x_3. \tag{63}$$

The potential $\mathcal{V}(c)$, which differs for the spherical geometry from the one in the torus case [12], is both gauge and rotational invariant and can be expressed entirely in terms of the x_i coordinates,

$$\frac{1}{2\pi^2}\mathcal{V}(c) = 2c_i^a c_i^a + 6\det c + \tfrac{1}{4}\left\{(c_i^a c_i^a)^2 - (c_i^a c_j^a)^2\right\}$$
$$= 2x_i^2 + 6\prod x_i + \tfrac{1}{2}\sum_{i>j} x_i^2 x_j^2. \tag{64}$$

Expanding around the sphaleron path we write

$$c_i^a = -u\delta_i^a + \tilde{c}_i^a, \quad \mathrm{tr}(\tilde{c}) \equiv \sum_i \tilde{c}_i^i = 0, \tag{65}$$

and first follow the steps relevant to integrating out $\tilde{c}$ in a quadratic approximation. This can be immediately taken from the analysis in sect. 2, by restricting eqs. (19), (23) and (24) to $\boldsymbol{L}_1^2 = 0$. The frequencies of the perpendicular $\tilde{c}$ fluctuations are thus read off from eqs. (A.5) and (A.6). They are labelled by k, the spin quantum number for $\boldsymbol{K} \equiv \boldsymbol{T} + \boldsymbol{S}$,

$$k = 1: \quad \omega_1^2 = 4u^2 - 6u + 4 \quad \text{(triplet)},$$
$$k = 2: \quad \omega_2^2 = 6u + 4 \quad \text{(quintet)}. \tag{66}$$

This overcounts the number of degrees of freedom, which is compensated by the ghosts. For $\tilde{c} = 0$ one easily deduces

$$\mathcal{M}_{\rm gh} = -D_i^2 = (2\boldsymbol{L}_1 + u\boldsymbol{T})^2. \tag{67}$$

Note that its spectrum is easily found in explicit form when writing

$$\mathcal{M}_{\rm gh} = 2(2-u)\boldsymbol{L}_1^2 + 2u(\boldsymbol{L}_1 + \boldsymbol{T})^2 - 2u(2-u), \tag{68}$$

and is positive definite everywhere, except at $u = 0$ (the $L_1^2 = 0$ mode) and at $u = 2$ (the $(\boldsymbol{L}_1 + \boldsymbol{T})^2 = 0$ mode), which are associated with the perturbative vacua. It is yet another way to demonstrate that these are singular points, as discussed extensively before. The quadratic approximation for integrating out the modes perpendicular to the sphaleron path, is well defined for all modes with non-zero zero-point frequencies. Thus, problems can only arise in the ghost sector near the classical vacua. For example, restricted to $L_1^2 = 0$ we have three gauge modes and

$$\omega_{\rm gh}^2 = u^2 \qquad \text{(triplet)}. \tag{69}$$

The adiabatic quadratic approximation will fail at $u = 0$, but is expected to be valid for large values of u. In that case one obtains an effective potential by summing all zero-point frequencies

$$V_{\rm eff} = \tfrac{3}{2}\omega_1 + \tfrac{5}{2}\omega_2 - 3\omega_{\rm gh} = \tfrac{5}{2}\sqrt{6u+4} + 3\sqrt{u^2 - \tfrac{3}{2}u + 1} - 3|u|. \tag{70}$$

These results could of course be derived directly from expanding eq. (64) in terms of the $\tilde{c}$ modes around the sphaleron path,

$$\mathcal{V}(c) = 3\pi^2 u^2 (2-u)^2 + \pi^2(4 + 2u^2)\operatorname{tr}(\tilde{c}\tilde{c}^T) + \pi^2(6u - 2u^2)\operatorname{tr}(\tilde{c}^2) + O(\tilde{c}^3). \tag{71}$$

We decompose $\tilde{c}$ in a triplet antisymmetric part ($\tilde{c}^{(a)}$) and in a quintet traceless symmetric part ($\tilde{c}^{(s)}$). As usual the norm of a gauge field is given by $\| A \|^2 = -\int_{S^3} \operatorname{Tr}(A_i^2)$. We find

$$\mathcal{V}(c) = 3\pi^2 u^2 (2-u)^2 + \tfrac{1}{2}\omega_1^2 \| \tilde{c}^{(a)} \|^2 + \tfrac{1}{2}\omega_2^2 \| \tilde{c}^{(s)} \|^2. \tag{72}$$

Also the kinetic term is diagonal, since

$$\| \dot{c} \|^2 = 3\pi^2 \dot{u}^2 + \| \dot{\tilde{c}}^{(a)} \|^2 + \| \dot{\tilde{c}}^{(s)} \|^2. \tag{73}$$

From these two equations one easily reproduces eq. (70).

The decomposition of eq. (60) allows us to rigorously eliminate the gauge degrees of freedom, which we will only address in the sector of constant modes. Gauge invariance is imposed by putting $\hat{T}^2 = 0$ and the angular degrees of freedom are integrated out by putting $\hat{L}^2 = 0$. This reduces the problem to three dimensions, but we now see that the sphaleron mode, which has $x_1 = x_2 = x_3$, corresponds to the most singular configuration. There, the jacobian $\mathcal{J}$ has a zero of third order, see eq. (62). It is however a coordinate singularity and can be removed by scaling the wave function by $\sqrt{\mathcal{J}}$. The rescaled wave function has to vanish as $\sqrt{\mathcal{J}}$, in particular at the sphaleron path (a similar situation arises in the torus geometry [8,12]). After this rescaling we find for the reduced hamiltonian

$$\hat{\mathcal{H}} = -\frac{g^2}{4\pi^2}\frac{\partial^2}{\partial x_i^2} - \frac{g^2}{4\pi^2}\sum_{i>j}\frac{x_i^2 + x_j^2}{(x_i^2 - x_j^2)^2} + \frac{\mathcal{V}(x)}{g^2}. \tag{74}$$

We now decompose x_i in terms of the sphaleron and perpendicular degrees of freedom

$$x_i = -u + \alpha_i y_1 + \beta_i y_2, \qquad \sum \alpha_i = \sum \beta_i = \sum \alpha_i \beta_i = 0, \qquad \sum \alpha_i^2 = \sum \beta_i^2 = 1, \tag{75}$$

with restrictions on y_1 and y_2 implied by eq. (63). These can also be read off from

$$\mathcal{J} = \prod_{i>j} |(\alpha_j - \alpha_i)y_1 + (\beta_j - \beta_i)y_2| \prod_k |2u + \alpha_k y_1 + \beta_k y_2|. \tag{76}$$

We expand the potential $\mathcal{V}(x)$ to second order in y_i

$$\mathcal{V}(x) = 3\pi^2 u^2 (2-u)^2 + \pi^2 \omega_1^2 (y_1^2 + y_2^2) + \mathrm{O}(y^3). \tag{77}$$

The fact that the reduced wave function is required to vanish for $\alpha_k y_1 + \beta_k y_2 = -2u$ will only have exponentially small effects (proportional to $\exp(-\gamma\sqrt{u})$ for some constant γ) and can be ignored for large u. To leading order in ω_1, integrating out the y_i modes is achieved by computing the ground state for the two-dimensional hamiltonian*

$$\hat{\mathcal{H}}(\hat{y}) = \tfrac{1}{2}\omega_1 \left\{ -\frac{\partial^2}{\partial \hat{y}_i^2} + (\hat{y}_1^2 + \hat{y}_2^2) - \frac{9(\hat{y}_1^2 + \hat{y}_2^2)^2}{4\hat{y}_1^2(\hat{y}_1^2 - 3\hat{y}_2^2)^2} \right\}, \tag{78}$$

with a wave function that vanishes as $\sqrt{\hat{\mathcal{J}}}$ where

$$\hat{\mathcal{J}} = \hat{y}_1(\hat{y}_1^2 - 3\hat{y}_2^2), \qquad \hat{y}_i = \frac{\pi\sqrt{2\omega_1}}{g} y_i. \tag{79}$$

Remarkably, a direct computation shows that $\hat{\Psi}(\hat{y}) = \sqrt{\hat{\mathcal{J}}}\exp(-\tfrac{1}{2}\hat{y}_i^2)$ is the ground state with an energy $\tfrac{5}{2}\omega_1$. In leading order this, as it should, coincides with eq. (70).

In principle, integrating out the y_i modes is well defined, but even if the adiabatic approximation holds all the way down to $u = 0$ and $u = 2$, it is very difficult to reliably determine the effective potential, due to the complicated structure of eq. (76). It would give us a one-dimensional effective hamiltonian on the interval $[0,2]$, which when periodically extended allows a straightforward implication of θ-dependence. Care is however required in extending the effective wave function to $u = 0$ and 2, due to the rescaling with $\sqrt{\mathcal{J}}$. We will not pursue this any further as it would only allow us to compute the ground state energy in each θ-sector. That no glueball excitations are expected to be reliably included, can be seen by considering the spectrum in lowest order in perturbation theory. It can be read off from eq. (24) at $u = 0$,

$$\mathcal{M}(u=0) = 2\boldsymbol{L}_2^2 + 2(\boldsymbol{L}_1 + \boldsymbol{S})^2. \tag{80}$$

* For definiteness we choose $\alpha_k = \sqrt{\tfrac{2}{3}}\sin(\tfrac{2}{3}\pi k)$ and $\beta_k = \sqrt{\tfrac{2}{3}}\cos(\tfrac{2}{3}\pi k)$.

The sphaleron mode has $\mathcal{M} = 4$ ($\boldsymbol{L}_1^2 = \boldsymbol{L}_2^2 = 0$). However, it is not the lowest mode, which is obtained by taking $\boldsymbol{L}_1^2 = \boldsymbol{L}_2^2 = (\boldsymbol{L}_1 + \boldsymbol{S})^2 = \frac{3}{4}$. These 12 transverse ($\partial_i A_i = 0$) eigenmodes are described in terms of three constant four-vectors $w_\mu^{(a)}$ by

$$\boldsymbol{A} = i \sum_{a=1}^{3} \frac{\tau_a}{2} \mathrm{Tr}\left(\frac{\tau_a}{2} \left\{ w^{(a)} \cdot n\boldsymbol{\tau} + \boldsymbol{w}^{(a)} \wedge \boldsymbol{\tau} \right\} \right) , \tag{81}$$

which coincides with the ε-mode at the sphaleron, if $w_\mu^{(k)} = -\varepsilon_\mu$ (cf. eqs. (A.14) and (A.15)).

When we are interested in the glueball spectrum, all perturbative modes should be treated at more equal footing. The sphaleron configurations are the ones that should be most sensitive to the θ-dependence. Sect. 6 will show how this can be formulated without affecting the perturbative results.

6. Imposing boundary conditions

The nine constant modes with $\boldsymbol{L}_1^2 = 0$ are degenerate with the nine modes that satisfy $\boldsymbol{L}_2^2 = 2$ and $\boldsymbol{L}_1 + \boldsymbol{S} = \boldsymbol{0}$. These 18 perturbatively degenerate modes with $\mathcal{M} = 4$ will feature prominently in the following, for reasons that will become clear shortly. The additional nine modes are easily seen to be given by

$$A_i = -\frac{i}{2} \hat{c}_i^a \tau_a, \quad \hat{c}_i^a = V_i^b d_b^a , \tag{82}$$

where d_b^a is constant and $V \in \mathrm{SO}(3)$ (an S^3-dependent rotation of the tangent frame)

$$V_i^b = \tfrac{1}{2} \mathrm{Tr}(n \cdot \bar{\sigma} \sigma_i n \cdot \sigma \sigma_b) . \tag{83}$$

Using

$$\sum_{\mu=1}^{4} \sigma_\mu^{\alpha\beta} (\bar{\sigma}_\mu)^{\gamma\delta} = 2\delta_{\alpha\delta}\delta_{\beta\gamma} , \tag{84}$$

it is easily verified that indeed $V \in \mathrm{SO}(3)$. One can also readily check that $\boldsymbol{L}_1 + \boldsymbol{S}$ vanishes on these modes, using the following identity (proven in appendix D):

$$L_1^a n \cdot \bar{\sigma} \tau_b n \cdot \sigma = -S_{bc}^a n \cdot \bar{\sigma} \tau_c n \cdot \sigma . \tag{85}$$

Note that all 18 modes are transverse

$$\partial_i c_i^a = \partial_i \hat{c}_i^a = 0 . \tag{86}$$

Crucial is now to observe that the modes c and $\hat{c}$ we have isolated, do not only contain the sphaleron mode $A_i = -iu\tau_i/2$, but also the mode related to this by the gauge transformation with winding number -1, $\Omega = n\cdot\overline{\sigma}$ (compare the discussion between eqs. (11) and (12))

$$n\cdot\overline{\sigma}(-iu\frac{\tau_i}{2})n\cdot\sigma + n\cdot\overline{\sigma}\partial_i n\cdot\sigma = -n\cdot\overline{\sigma}\left(-i(2-u)\frac{\tau_i}{2}\right)n\cdot\sigma$$
$$= -i(2-u)V_i^j\frac{\tau_j}{2}. \qquad (87)$$

This gauge transformation maps $c_i^a = -u\delta_i^a$ to $d_i^a = (u-2)\delta_i^a$. Thus the 18 modes contain three perturbative vacua and two sphaleron modes

$$c_i^a = -u\delta_i^a, \quad d_i^a = -v\delta_i^a. \qquad (88)$$

The sphalerons correspond to respectively $u = 1$ and $v = 1$. This will be confirmed by computing the classical potential for $A_i = i(c_i^a - \widehat{c}_i^a)\tau_a/2$.

As a first step to compute the potential $\mathcal{V}(A)$, we evaluate the magnetic field

$$B_k = \tfrac{1}{2}\varepsilon_{ijk}F_{ij} = B_k(c) + V_k^l B_l(d) + \widehat{B}_k(c,d), \qquad (89)$$

with (cf. eq. (13))

$$B_k(c) \equiv -2c_k + \tfrac{1}{2}\varepsilon_{ijk}[c_i,c_j], \qquad c_i \equiv ic_i^a\frac{\tau_a}{2},$$
$$B_k(d) \equiv -2d_k + \tfrac{1}{2}\varepsilon_{ijk}[d_i,d_j], \qquad d_i \equiv id_i^a\frac{\tau_a}{2},$$
$$\widehat{B}_k(c,d) \equiv -\varepsilon_{ijk}[c_i,\widehat{c}_j] = -\varepsilon_{ijk}V_j^l[c_i,d_l], \qquad \widehat{c}_i \equiv i\widehat{c}_i^a\frac{\tau_a}{2}. \qquad (90)$$

To compute the potential, the non-trivial integrals over S^3 will involve

$$\int_{S^3} V_a^b = 0, \quad \int_{S^3} V_a^b V_c^d = \frac{2\pi^2}{3}\delta_{ac}\delta_{bd}. \qquad (91)$$

The second identity requires a proof, which is deferred to appendix D. One finds now

$$\mathcal{V}(c,d) \equiv -\int_{S^3}\mathrm{Tr}\left(B_k^2\right) = \mathcal{V}(c) + \mathcal{V}(d) + \frac{2\pi^2}{3}\left\{(c_i^a)^2(d_j^b)^2 - (c_i^a d_j^a)^2\right\} \qquad (92)$$

(cf. eq. (64) for the definition of $\mathcal{V}(c)$). Observe that the potential has an SO(3)×SO(3) rotational invariance. This is not accidental. It is a consequence of the fact that $\boldsymbol{L}_2$ acts on $n\cdot\overline{\sigma}\tau_b n\cdot\sigma$ as a gauge transformation (see appendix D for a derivation)

$$L_2^a n\cdot\overline{\sigma}\tau_b n\cdot\sigma = -T^a n\cdot\overline{\sigma}\tau_b n\cdot\sigma, \qquad (93)$$

from which it is easily verified that $\boldsymbol{L}_2$ acts on d as a rotation,

$$L_2^a\widehat{c}_i = -V_i^k(S^a d)_k. \qquad (94)$$

Generally, the potential will rapidly increase in all directions except for the directions of the two sphalerons. It is only in the latter directions that we wish to implement, beyond the semiclassical approximation, the effect of a non-zero value of θ. We do this here through appropriate boundary conditions in configuration space. This is not too different from what was done in the torus geometry [8,18]. A more general point of view was advocated in ref. [20]. As we have seen, the various perturbative vacua are mapped onto each other by gauge transformations with non-zero winding number. The hamiltonian in the Coulomb gauge [35] is regular at $A = 0$. It can be extended up to the Gribov horizon, where the vanishing of the Faddeev–Popov determinant $\det'(\mathcal{M}_{\rm gh})$

$$\mathcal{M}_{\rm gh} \equiv -\partial_i D_i(A) \tag{95}$$

(the prime indicates that the determinant is evaluated on the subspace with $L_1^2 \neq 0$) makes it singular. The perturbative vacua defined by $u = 2$ or $v = 2$ (all other modes vanishing) have vanishing Faddeev–Popov determinants. Around each, a hamiltonian is defined by conjugating the hamiltonian around $A = 0$ with the gauge function that maps these vacua to $A = 0$. Where configurations overlap, the transition function is precisely this gauge transformation. This construction was discussed in detail in ref. [18].

Alternatively [20] one imposes boundary conditions on the boundary $\partial\Lambda$ of the fundamental modular domain Λ. The fundamental modular domain is the convex set of transverse gauge potential, which are such that the function I_A defined on the space of local gauge transformations has its absolute minimum at the unit gauge function. This function I_A is simply the norm of the gauge transformed vector potential

$$I_A(\Omega) = I_{[\Omega]A}(g) = \| [\Omega]A \|^2 = -\int_{S^3} \mathrm{Tr}(A_i - \Omega^{-1}\partial_i\Omega)^2, \quad g \in \mathrm{G}. \tag{96}$$

We remind the reader of the fact that stationarity of this functional implies the Coulomb gauge condition and that the hessian is given by eq. (95). The boundary of the fundamental modular domain is defined by those transverse configurations where the absolute minimum at* $\Omega = g$ is degenerate with an absolute minimum at $\Omega_0 \neq g$. The other minimum corresponds to another point, $[\Omega_0]A$, on $\partial\Lambda$. Obviously, the two points on $\partial\Lambda$ are related by the gauge transformation Ω_0. Under this identification the wave functional picks up a phase $\exp(in\theta)$ where n is the winding number of the gauge function Ω_0,

$$n(\Omega) = \frac{1}{24\pi^2}\int_{S^3} \mathrm{Tr}(\Omega^{-1}\mathrm{d}\Omega)^3 . \tag{97}$$

One explicitly verifies that the two sphalerons,

$$A_j^{(1)} = -i\frac{\tau_j}{2}, \quad A_j^{(2)} = in\cdot\bar{\sigma}\frac{\tau_j}{2}n\cdot\sigma , \tag{98}$$

* As in eq. (96), g denotes an arbitrary *constant* gauge function.

are conjugate points on $\partial\Lambda$, as $A^{(2)} = [n \cdot \overline{\sigma}]A^{(1)}$ and $I_A(g)$ is identical for both fields. As c and $\widehat{c}$ are orthogonal, one even has

$$I_{c-\widehat{c}}(1) = \|c\|^2 + \|\widehat{c}\|^2 = \|c\|^2 + \|d\|^2. \tag{99}$$

A useful further ingredient to study boundary identifications is the Chern–Simons functional

$$Q(A) = \frac{1}{8\pi^2} \int_{S^3} \mathrm{Tr}(A_i B_i - \tfrac{1}{3}\varepsilon_{ijk} A_i A_j A_k). \tag{100}$$

One easily shows that along the tunnelling path [eq. (9)] $Q(A(s,b))$ coincides with $Q(s,b)$, as given in eq. (35). Furthermore

$$Q(c) = \tfrac{1}{4}(c_i^a c_i^a + \det c), \quad Q(c - \widehat{c}) = Q(c) - Q(d). \tag{101}$$

Note that the perturbative vacua have integer and the sphalerons half-integer Chern–Simons values and that

$$Q([\Omega]A) = Q(A) + n(\Omega). \tag{102}$$

In the following H_λ denotes the subspace of modes with an eigenvalue λ for $\mathcal{M}(u = 0)$ (eq. (80)) and $\widehat{\mathrm{H}}_\lambda$ denotes its restriction to the transverse modes $\partial_j A_j = -2iL_1^j A_j = 0$. As $[L_1^j, \mathcal{M}(u = 0)] = 0$, this projection is well defined. We have seen earlier that $\widehat{\mathrm{H}}_3 = \mathrm{H}_3$ and $\widehat{\mathrm{H}}_4 = \mathrm{H}_4$. As long as the energies remain comparable to the sphaleron energy, the wave functional will decay rapidly in all directions of configuration space $\oplus_\lambda \widehat{\mathrm{H}}_\lambda$, except in the directions of the two sphaleron modes that connect the two perturbative vacua nearest to $A = 0$. These sphaleron modes are contained in $\widehat{\mathrm{H}}_4$. We therefore take all the modes in $\widehat{\mathrm{H}}_\lambda$ for $\lambda \neq 4$ gaussian, ignoring the boundary conditions at $\widehat{\mathrm{H}}_\lambda \bigcap \partial\Lambda$. Boundary conditions will be formulated exclusively in $\widehat{\mathrm{H}}_4 \bigcap \partial\Lambda$. However, only *at* the sphalerons, the boundary map identifies points within $\widehat{\mathrm{H}}_4$ (namely mapping one sphaleron into the other). To see this (the various symmetries allow us to remove gauge and rotational degrees of freedom) take $c_i = ix_i\tau_i/2$. Applying the gauge transformation $\Omega = n \cdot \overline{\sigma}$ yields $A_i = in \cdot \overline{\sigma}(x_i - 2)\tau_i n \cdot \sigma/2$. It still has $\boldsymbol{L}_1^2 = \boldsymbol{L}_2^2 = 2$, but it is easily seen that $(\boldsymbol{L}_1 + \boldsymbol{S})^2 A_i = 0$, iff $x_1 = x_2 = x_3$. Nevertheless, also on $\widehat{\mathrm{H}}_4$ the wave functional decays rapidly in all directions away from the sphalerons. Hence, except for the sphalerons, the points on $\partial\Lambda$ couple the various subspaces $\widehat{\mathrm{H}}_\lambda$ under the boundary identification, but in all these directions boundary conditions are irrelevant. It implies, that to a good approximation, the wave functional near the sphaleron $A^{(1)}$ can be decomposed as

$$\Psi(A) = \varphi_1(u)\chi_{[u]}(P_1 A), \tag{103}$$

and near the sphaleron $A^{(2)}$ as

$$\Psi(A) = \varphi_2(v)\chi_{[v]}(P_2 A), \tag{104}$$

where (see eq. (98))

$$u \equiv -\frac{1}{3\pi^2}\int_{S^3} \mathrm{Tr}(A_j A_j^{(1)}), \quad v = -\frac{1}{3\pi^2}\int_{S^3} \mathrm{Tr}(A_j A_j^{(2)}) \tag{105}$$

are the components of A along the sphaleron modes, whereas

$$P_i(A) = A + \frac{A^{(i)}}{3\pi^2}\int_{S^3} \mathrm{Tr}(A_j A_j^{(i)}) \tag{106}$$

projects onto the directions perpendicular to these modes. The θ-dependence is implemented by imposing

$$\varphi_1(1) = \mathrm{e}^{i\theta}\varphi_2(1). \tag{107}$$

If χ factorises on $\widehat{\mathrm{H}}_4 \oplus \widehat{\mathrm{H}}_4^{\perp}$, the boundary condition can be formulated entirely within $\widehat{\mathrm{H}}_4$. We will call this condition the sphaleron factorisation property. It is very likely to be justified as *near* the sphalerons the time scale for fluctuations along the sphaleron path is much longer than the time scale associated to the perpendicular fluctuations. Much like in the torus case [8], one can in principle check a posteriori in how far this adiabatic decomposition is valid at the sphalerons.

Under the sphaleron factorisation hypothesis one is insensitive to replacing $\partial\Lambda \bigcap \widehat{\mathrm{H}}_4$ by $\partial\Lambda_4$, provided both coincide near, and at, the sphalerons and $\partial\Lambda_4$ encloses $\partial\Lambda \bigcap \widehat{\mathrm{H}}_4$ (as otherwise the wave functional at $A \in \partial\Lambda_4$ could be appreciable, whereas it is negligible at $\{\mu A | \mu \in \mathbb{R}^+\} \bigcap \partial\Lambda$). Points near $A^{(1)}$ are to be identified with points near $A^{(2)}$ on $\partial\Lambda_4$. One knows quite a lot about $\partial\Lambda$. We already mentioned that Λ is convex [20,37], but furthermore, it was shown that in each direction in configuration space the Gribov horizon, and therefore also the boundary of the fundamental modular domain, is at a finite distance from the origin [38]. Hence $\partial\Lambda \bigcap \widehat{\mathrm{H}}_\lambda$ is compact. It is in itself an interesting problem to find $\partial\Lambda$ restricted to, say, $\widehat{\mathrm{H}}_3 \oplus \widehat{\mathrm{H}}_4$. But under the sphaleron factorisation hypothesis a precise knowledge is not really required. All we wish to state here is that one might suspect that $\partial\Lambda \bigcap \widehat{\mathrm{H}}_4$ is given by $|Q(c-\hat{c})| = \frac{1}{2}$. This is false, as Q vanishes along the line $c = d$; a line which intersects $\partial\Lambda$.

In general, we have to construct $\partial\Lambda_4$ such that it is preserved under rotations and constant gauge transformations. It is most suitable to have it also invariant under interchanging c and d, in which case we can implement the boundary condition by

$$\Psi(c,d) = \mathrm{e}^{i\theta}\Psi(d,c), \quad (c,d) \in \partial\Lambda_4^{(1)}. \tag{108}$$

As is indicated in fig. 2, we take two disconnected branches for $\partial\Lambda_4$, this is to ensure that the closed loop, that occurs by identifying two points on the boundary, will be non-contractable. Otherwise θ would not be a good quantum number. In particular, $\partial\Lambda_4$ should avoid the line where $c = d$, else it will

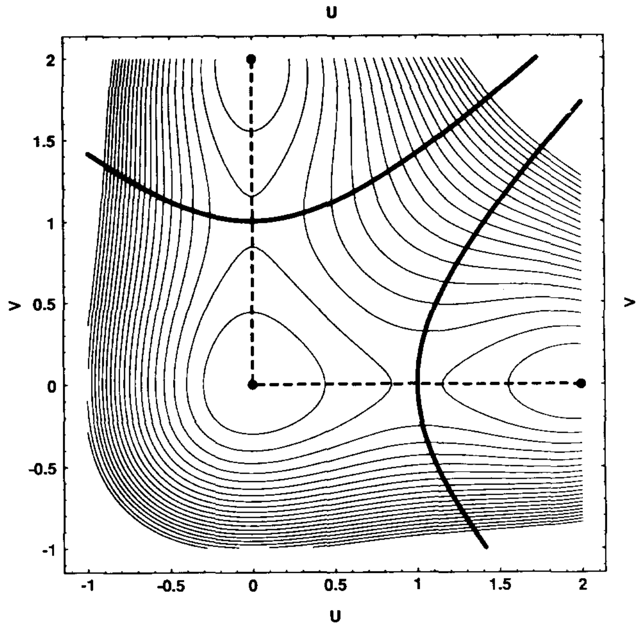

Fig. 2. The equipotential lines of the classical potential (see eq. (92)), restricted to the sphaleron modes of eq. (88). The outermost contour corresponds to an energy roughly ten times the sphaleron energy. The classical vacua are indicated by the large dots, the tunnelling paths through the sphalerons by the dashed lines. Boundary conditions (eq. (108)) are implemented on the fat curves (eq. (109)), which near the spahalerons coincide with the boundary of the fundamental modular domain.

force the wave function (rigorously) to zero there. In fig. 2 we have also plotted the equipotential lines up to an energy ten times that of the sphaleron, using the potential of eq. (92), restricted to the sphaleron modes of eq. (88). Observe the very steep rise of the potential in the direction perpendicular to the sphaleron paths. Just to give an explicit example, fig. 2 is based on the choice

$$\partial \Lambda_4^{(i)} = \{(c,d)|(\|d\|^2 - \|c\|^2) = 3(-1)^i\pi^2\}. \tag{109}$$

such that $\partial \Lambda_4$ intersects the sphaleron path perpendicularly*. Note that it satisfies the crucial requirement, that gauge and rotational invariance are not broken by the boundary conditions. This completes our discussion of how we implement the θ-dependence. Sect. 7 explains how this should be incorporated in future calculations of the glueball spectrum.

* Part of the true boundary of the fundamental modular domain, restricted to the space represented in fig. 2, as well as the Gribov horizon, were recently constructed in ref. [43].

7. Discussion and conclusions

In sect. 6 we have introduced our proposal for including the θ-dependence on the glueball spectrum. It is implemented as follows: in perturbation theory, to lowest order, all modes are quadratic and the spectrum is given by $\widehat{\mathcal{M}} \equiv \mathcal{M}(u = 0) = 2\boldsymbol{L}_2^2 + 2(\boldsymbol{L}_1 + \boldsymbol{S})^2$, with eigenvalues λ (their precise values are presently irrelevant). Using the Yang–Mills hamiltonian in the coulomb gauge [35], the physical Hilbert space is given by the transverse modes $\oplus_\lambda \widehat{\mathrm{H}}_\lambda$, where $\widehat{\mathrm{H}}_\lambda$ is the n_λ-dimensional subspace of transverse fields with eigenvalue λ. The hamiltonian is given by [6,7,35]

$$\mathcal{H} = \tfrac{1}{2}g_0^2 \int\int_{S^3} \rho^{-1/2}\pi_i^a(x)\rho^{1/2}K(x,y)_{ij}^{ab}\rho^{1/2}\pi_j^b(y)\rho^{-1/2} + \frac{1}{2g_0^2}\int_{S^3}(B_k^a(x))^2 . \tag{110}$$

The momenta of the transverse fields are expressed in terms of the electric field (which is canonically conjugate to the gauge field) as

$$\pi_k^a(x) = E_k^a(x) - \partial_k \Delta^{-1}\partial_l E_l^a(x) . \tag{111}$$

Furthermore, Δ^{-1} denotes the Green function for the laplacian on S^3, $\rho = \det'(-\partial_k D_k)$ is the Faddeev–Popov determinant and $K(x,y)$ is the kernel

$$K(x,y)_{ij} = \delta_{ij}\delta_3(x-y) + \mathrm{ad}A_i(\partial_l D_l)^{-1}\Delta(\partial_m D_m)^{-1}\mathrm{ad}A_j . \tag{112}$$

In leading order $\mathcal{H}$ is given by

$$\mathcal{H}_0 = \tfrac{1}{2}g_0^2 \int_{S^3} \pi_k^a(x)^2 + \frac{1}{2g_0^2}\int_{S^3} A_i^a(x)\widehat{\mathcal{M}}_{ij}^{ab}A_j^b(x) . \tag{113}$$

Standard perturbation theory is performed in terms of the Fock space that consists of the occupation numbers for each oscillator. Our proposal to implement the θ-dependence is to replace in $\widehat{\mathrm{H}}_4$ the harmonic oscillator wave functions by the eigenfunctions that satisfy the θ-dependent boundary conditions of eq. (108). Details will be worked out in the future. The energies of the modes in $\widehat{\mathrm{H}}_4$ are no longer proportional to the particle number and matrix elements that involve field components in $\widehat{\mathrm{H}}_4$ will be more complicated. However, only a few modes suffer this complication and perturbation theory is rigorously embedded, as the effects of the boundary conditions are suppressed exponentially in the inverse coupling constant. There are no obstacles to formulate dimensional regularisation. One adds the ε extra dimensions, for example, in the form of a symmetric torus T^ε of size R. Nevertheless, one needs to demonstrate that all infinities can be appropriately absorbed in coupling constant and field renormalisations. As the ultraviolet divergencies are due to the high-energy modes, which are rigorously treated as in perturbation theory, this renormalisation should cause no major problems, provided we formulate the boundary conditions for $\widehat{\mathrm{H}}_4$ in terms of the renormalised fields.

The present proposal is well defined, without ambiguities. It preserves all symmetries, most importantly including the topological symmetry, which is crucial in the domain where θ-dependence becomes appreciable. In sect. 1 we have already indicated why the rewards of this endeavour might be worth its price. Ultimately, of course, the aim would be to take the infinite-volume limit, and above all understand the origin of colour confinement. We believe this is presently still too ambitious, but we have learned that the boundary of the fundamental modular domain dominates to a large extend the dynamics for the low-lying spectrum.

We should mention the alternative approaches to include the non-trivial issues associated to the presence of Gribov copies. Closely related to the methods followed here, Vohwinkel is attempting to address this in the torus geometry, by analysing the contribution of the first few higher momentum modes [39]. Other attempts that have been under development for quite a few years now can be found in the works of Cutkosky and collaborators [7]. Like in the present paper they use the hamiltonian formulation in a spherical geometry. In ref. [40] an account of the present status is given. Zwanziger and collaborators [38] are approaching the problem from the euclidean path integral end, largely in the context of lattice gauge theory. A recent paper incorporates the restriction to the fundamental modular domain and contains a review of most of the earlier results [41].

In conclusion, this paper has explored ways of incorporating θ-dependence in the glueball spectrum and the vacuum energy, that essentially go beyond the semiclassical approximation. In principle, although rather impractical, vacuum energies can be obtained from expanding around the (single) tunnelling path that goes through the mountain pass (sphaleron) separating two valleys (classical vacua). Glueballs can only be included by treating perturbative modes more or less at the same footing. But 18 modes, degenerate in lowest quadratic order (but not of lowest energy) are singled out for implementing the θ-dependence through boundary conditions, derived from the properties of the boundary of the Coulomb-gauge fundamental modular domain. Time will be the judge of the feasibility of the proposed programme. We herewith rest our case.

N.D.H. thanks the ITF of the University of Amsterdam for the invitation to spend a year during which this work was done. P.v.B. thanks Valentin Khoze, Martin Lüscher, Andreas Ringwald, Jac Verbaarschot and Christof Wetterich for discussions and Dick Cutkosky and Dan Zwanziger for correspondence. In particular he wishes to thank Claus Vohwinkel for sharing his insights on attempts to include higher momentum modes in the torus geometry. He is also grateful for Peter Braam's early help in formulating the rudiments of the moduli space approach. Hospitality during July 1991 at CERN, where some of this work was done is greatly appreciated. He also thanks Martin Lüscher and J. Bartels for hospitality at Desy in December 1991. The help of Mathematica [42] in performing some tedious algebra and producing the figures was appreciated. This research has been supported by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

Appendix A

Here we prove that the fluctuation operator $\mathcal{M}$, defined in eq. (19), is positive definite on all states perpendicular to the sphaleron mode σ_i. We repeat here eqs. (23) and (24), using $u = 1 + a$ and add a third way to write eq. (19):

$$\mathcal{M} = 2\Big\{(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 + a(\boldsymbol{L}_1 + \boldsymbol{T})^2 + (1-a)\boldsymbol{L}_1^2 + a(1-a)\boldsymbol{T}\cdot\boldsymbol{S} - (1-a^2)\Big\}, \tag{A.1}$$

$$\mathcal{M} = 2\Big\{(a+1)(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 - a(\boldsymbol{L}_1 + \boldsymbol{S})^2 + \boldsymbol{L}_1^2 - a(a+1)\boldsymbol{T}\cdot\boldsymbol{S} - (1-a^2)\Big\}, \tag{A.2}$$

$$\mathcal{M} = a(2\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 + 2\Big\{(1-a)(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2 + (1-a)\boldsymbol{L}_1^2 + a(1-a)\boldsymbol{T}\cdot\boldsymbol{S} + a^2 + a - 1\Big\}. \tag{A.3}$$

Since $\mathcal{M}$ commutes with $\boldsymbol{L}_1^2 = \boldsymbol{L}_2^2$, we can study the spectrum restricted to each value of l_1, where $\boldsymbol{L}_1^2 = l_1(l_1+1)$.

For $l_1 \geqslant 2$, eq. (A.1) shows that $\mathcal{M} \geqslant 4l_1(l_1 - 1 - a) + 6a^2 - 4a + 2 \geqslant \frac{4}{3}$ at $1 \geqslant a \geqslant 0$, whereas for $-1 \leqslant a \leqslant 0$, eq. (A.2) allows us to bound $\mathcal{M}$ by $6a^2 + 2 + 4l_1(l_1 - 1) - a(4l_1 - 8) \geqslant 2$. For $l_1 = 0$ we have

$$\mathcal{M}(l_1 = 0) = 6a^2 - 2 + (2 + a - a^2)(\boldsymbol{T} + \boldsymbol{S})^2, \tag{A.4}$$

which contains the sphaleron mode at $\boldsymbol{K} \equiv \boldsymbol{T} + \boldsymbol{S} = \boldsymbol{0}$,

$$\mathcal{M}(l_1 = k = 0) = 6a^2 - 2, \tag{A.5}$$

whereas the $k = 1$ triplet and $k = 2$ quintet have positive eigenvalues for $-1 \leqslant a \leqslant 1$:

$$\mathcal{M}(l_1 = 0, k = 1) = 4a^2 + 2a + 2, \qquad \mathcal{M}(l_1 = 0, k = 2) = 10 + 6a. \tag{A.6}$$

When $l_1 = \frac{1}{2}$ or $\frac{3}{2}$, the operators $(\boldsymbol{L}_1 + \boldsymbol{S} + \boldsymbol{T})^2$, $(\boldsymbol{L}_1 + \boldsymbol{T})^2$, $(\boldsymbol{L}_1 + \boldsymbol{S})^2$ and $\boldsymbol{L}_1^2$ are all bounded from below by $\frac{3}{4}$, whereas $\boldsymbol{T}\cdot\boldsymbol{S} \geqslant -2$. Using eq. (A.1) for $a \geqslant 0$ and eq. (A.2) for $a \leqslant 0$, one easily deduces $\mathcal{M} \geqslant 6a^2 - 4|a| + 1 \geqslant \frac{1}{3}$. Finally, only for $l_1 = 1$ eq. (A.3) will be required. For $a \geqslant 0$ it implies

$\mathcal{M} \geqslant 6a^2 - 6a + 2 \geqslant \frac{1}{2}$. For $-1 \leqslant a \leqslant 0$, we can use eq. (A.2), from which $\mathcal{M} \geqslant 6a^2 + 4a + 2 \geqslant \frac{4}{3}$. This completes the proof of the positivity of $\mathcal{M}$ for all modes perpendicular to the sphaleron mode and for $a^2 \leqslant 1$.

It is worthwhile to analyse a potentially disturbing problem associated with gauge fixing. When the background field does not satisfy the equations of motion, it seems superficially that longitudinal and transverse modes couple, which would destroy gauge invariance. This comes about when one splits the fluctuations in transverse and longitudinal modes. Their respective projection operators are P and $Q = 1 - P$ with

$$(Qq)_i = D_i D_l^{-2} D_j q_j \,. \tag{A.7}$$

One easily verifies that (with ϕ a Lie algebra valued function on S^3),

$$iD_j\phi = (2L_1^j + (1+a)T^j)\phi, \tag{A.8}$$

$$(\mathcal{M}D)_i = -D_i D_l^2 - \mathrm{ad}(D_j F_{ji}), \tag{A.9}$$

which implies that $P\mathcal{M}Q \neq 0$, since

$$(\mathcal{M}Q)_{ij} = -D_i D_j - \mathrm{ad}(D_k F_{ki}) D_l^{-2} D_j \,. \tag{A.10}$$

Only when $D_k F_{ki} = 0$ one has

$$\mathcal{M}_{ij} = (P\mathcal{M}P)_{ij} - D_i D_j \,. \tag{A.11}$$

The resolution of this apparent discrepancy lies in the fact that a variation in the direction of the gauge orbit contains a quadratic term in the gauge parameter that, combined with the term proportional to the non-vanishing equation of motion term, cancels against what comes from the seemingly offending second term in eq. (A.10). A gauge variation with $\Omega = \exp(-X)$ is given by

$$q_i = [\Omega]A_i - A_i = \mathrm{e}^{-X} D_i \mathrm{e}^{X} = D_i X + \tfrac{1}{2}[D_i X, X] + \mathrm{O}(X^3) \,. \tag{A.12}$$

Varying the potential $\mathcal{V}$ gives

$$\delta\mathcal{V} = \int_{S^3} 2\mathrm{Tr}(D_k F_{ki} q_i) - \mathrm{Tr}(q_i(\mathcal{M}_{ij} + D_i D_j) q_j) \,. \tag{A.13}$$

As $D_i(D_k F_{ij}) = 0$, the first term in eq. (A.13) contributes at second order $\int_{S^3} \mathrm{Tr}(D_k F_{ki}[D_i X, X])$, whereas the second term contributes, using eq. (A.9), $\int_{S^3} \mathrm{Tr}((D_i X)\mathrm{ad}(D_j F_{ji}) X)$. The two are easily seen to cancel.

We end this appendix with explicitly demonstrating the positivity of $\mathcal{M}$ in the ε-direction (thereby confirming that the seeming instability of $\mathcal{V}$ in this direction is caused by the non-linearities). The two-dimensional space spanned by ($\|q^{(1)}\| = \|q^{(2)}\|$ and $l_1 = \frac{1}{2}$)

$$\boldsymbol{q}^{(1)} = \sqrt{2}(\varepsilon \cdot n)\boldsymbol{\tau}, \qquad \boldsymbol{q}^{(2)} = (\boldsymbol{\varepsilon} \wedge \boldsymbol{\tau}), \tag{A.14}$$

is left invariant by $\mathcal{M}$. The reduced 2×2 matrix for $\mathcal{M}$ on this basis is

$$\mathcal{M}^{(ij)} = \begin{pmatrix} 6a^2+1 & 2a\sqrt{2} \\ 2a\sqrt{2} & 4a^2+1 \end{pmatrix} \tag{A.15}$$

with eigenvalues $\lambda = 1 + 5a^2 \pm \sqrt{a^2(a^2+8)}$. These are positive for all a.

Appendix B

The purpose of this appendix is to list the solution of the valley equation, eq. (37), when expanding A_i around $-iu\tau_i/2$ to third order in the four-vector ε_μ. Its expansion is given in eqs. (41) and (42). Up to its symmetries, the valley equation is solved by $(u = a + 1)$

$$f = -\tfrac{1}{2}(1+a) + \tfrac{1}{2}a(\varepsilon\cdot n) + f_0(\varepsilon\cdot n)^2 + f_1(\varepsilon\cdot n)^3, \tag{B.1}$$

$$g = -\tfrac{1}{2} + g_0(\varepsilon\cdot n) + g_1(\varepsilon\cdot n)^2, \tag{B.2}$$

$$h = 0, \tag{B.3}$$

$$\lambda_0 = -2a(1-a^2) + (4f_0 + 4ag_0 - 3a)(\varepsilon\cdot\varepsilon), \tag{B.4}$$

$$\lambda_\mu = (1+4a^2)\varepsilon_\mu + \left\{1 + \frac{4(af_0 + a^2g_0 + 3af_1 + (1+2a^2)g_1)}{1-a^2}\right\}(\varepsilon\cdot\varepsilon)\varepsilon_\mu, \tag{B.5}$$

$$\Lambda = \tfrac{5}{2}a + (2f_0 - \tfrac{3}{2}a)(\varepsilon\cdot n) + (a + 2f_0 + 2ag_0 + 3f_1)(\varepsilon\cdot n)^2 + \frac{2(f_0 + ag_0 + 3f_1 + 3ag_1)}{1-a^2}(\varepsilon\cdot\varepsilon). \tag{B.6}$$

The remaining four equations are first-order differential equations for f_0, g_0, f_1 and g_1,

$$2a(1-a^2)\frac{\mathrm{d}f_0}{\mathrm{d}a} + 2(2-a^2)f_0 + 8ag_0 - 3a^3 = 0, \tag{B.7}$$

$$2a(1-a^2)\frac{\mathrm{d}g_0}{\mathrm{d}a} + 2(2-3a^2)g_0 + 4af_0 + \tfrac{3}{2}(3a^2-1) = 0, \tag{B.8}$$

$$2a(1-a^2)\frac{\mathrm{d}f_1}{\mathrm{d}a} + 2(5-3a^2)f_1 + 10ag_1 + 4(2-3a^2)f_0 - 4ag_0 + a(2+a^2) = 0, \tag{B.9}$$

$$2a(1-a^2)\frac{\mathrm{d}g_1}{\mathrm{d}a} + 10(1-a^2)g_1 + 6af_1 + 6(1-a^2)g_0 + \tfrac{1}{2}(1-7a^2) = 0. \tag{B.10}$$

A considerable simplification is achieved by defining

$$F_0 = \frac{a^2 f_0}{\sqrt{1-a^2}}, \quad G_0 = \tfrac{1}{2}a^2 g_0\sqrt{1-a^2}, \quad F_1 = \frac{a^5 f_1}{1-a^2}, \quad G_1 = a^5 g_1. \tag{B.11}$$

We find

$$\frac{dF_0}{da} + \frac{8G_0}{(1-a^2)^2} - \frac{3a^4}{2(1-a^2)^{3/2}} = 0, \tag{B.12}$$

$$\frac{dG_0}{da} + F_0 - \frac{3a(1-3a^2)}{8\sqrt{1-a^2}} = 0, \tag{B.13}$$

$$\frac{dF_1}{da} + \frac{5G_1}{(1-a^2)^2} + \frac{2a^2(2-3a^2)F_0}{(1-a^2)^{3/2}} - \frac{4a^3G_0}{(1-a^2)^{5/2}} + \frac{a^5(2+a^2)}{2(1-a^2)^2} = 0, \tag{B.14}$$

$$\frac{dG_1}{da} + 3F_1 + \frac{6a^2G_0}{\sqrt{1-a^2}} + \frac{a^4(1-7a^2)}{4(1-a^2)} = 0, \tag{B.15}$$

which reduce to

$$\partial_a^2 G_i - \frac{\lambda_i G_i}{(1-a^2)^2} = V_i, \quad \lambda_0 = 8, \quad \lambda_1 = 15. \tag{B.16}$$

These are solved through the Legendre equation

$$(1-a^2)\partial_a^2 Q_\nu^\mu - 2a\partial_a Q_\nu^\mu + (\nu(\nu+1) - \mu^2(1-a^2)^{-1})Q_\nu^\mu = 0. \tag{B.17}$$

The homogeneous solutions are easily found to be

$$G_0^{(1)} = \frac{1+3a^2}{1-a^2}, \qquad G_0^{(2)} = \frac{a(3+a^2)}{3(1-a^2)},$$
$$G_1^{(1)} = \frac{a(1+a^2)}{(1-a^2)^{3/2}}, \qquad G_1^{(2)} = -\frac{(1+6a^2+a^4)}{(1-a^2)^{3/2}}. \tag{B.18}$$

The inhomogeneous solutions now follow from

$$G_i(a) = G_i^{(2)}(a)\int_0^a G_i^{(1)}(\hat{a})V_i(\hat{a})d\hat{a} - G_i^{(1)}(a)\int_0^a G_i^{(2)}(\hat{a})V_i(\hat{a})d\hat{a}. \tag{B.19}$$

One easily verifies that

$$V_0 = \frac{3(1-9a^2+2a^4)}{8(1-a^2)^{3/2}},$$
$$V_1 = \frac{a^3(-13+84a^2-53a^4)}{4(1-a^2)^2} + \frac{2a^2(5-6a^2)F_0}{(1-a^2)^{3/2}} + \frac{2a(-6+7a^2-3a^4)G_0}{(1-a^2)^{5/2}}. \tag{B.20}$$

Unlike for G_1, we can express G_0 in terms of elementary functions:

$$G_0 = -\frac{47+70a^2+3a^4}{24\sqrt{1-a^2}} - \frac{5a(3+a^2)\arcsin a}{4(1-a^2)}. \tag{B.21}$$

Appendix C

In this appendix we discuss in some more detail the computation of the riemannian metric and curvature. The metric $\dot{s}^2$ is defined by eqs. (27) and (28). The only tricky part in the calculations involves the Green function for the covariant laplacian. It acts on $D_i\dot{A}_i$, which for the ansatz of eq. (41) has the form $(u = a + 1)$

$$D_i\dot{A}_i = i\chi_1(\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}) + i\chi_2(\dot{\varepsilon}\cdot n)(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}) + i\chi_3\boldsymbol{\varepsilon}\cdot(\dot{\boldsymbol{\varepsilon}}\wedge\boldsymbol{\tau}) + i\chi_4(\boldsymbol{\varepsilon}\cdot\dot{\boldsymbol{\varepsilon}})(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}) + i\chi_5\dot{a}(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}), \tag{C.1}$$

with, to the relevant order in ε,

$$\begin{aligned}\chi_1 &= \tfrac{3}{2}a + (-a + 2f_0 - 2ag_0)(\varepsilon\cdot n) + (-2f_0 + 2ag_0 + 3f_1 - 2ag_1)(\varepsilon\cdot n)^2 + \mathrm{O}(\varepsilon^3),\\ \chi_2 &= (a + 2f_0 - 2ag_0) + (4f_0 + 6f_1 - 4ag_1)(\varepsilon\cdot n) + \mathrm{O}(\varepsilon^2),\\ \chi_3 &= -\tfrac{1}{2} + 2g_0(\varepsilon\cdot n) + \mathrm{O}(\varepsilon^2), \qquad \chi_4 = 0, \qquad \chi_5 = -\tfrac{1}{2} + \mathrm{O}(\varepsilon).\end{aligned} \tag{C.2}$$

As a direct computation shows, $D_l^2(D_i\dot{A}_i)$ is again of the form of eq. (C.1):

$$D_l^2(D_i\dot{A}_i) = i\eta_1(\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}) + i\eta_2(\dot{\varepsilon}\cdot n)(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}) + i\eta_3\boldsymbol{\varepsilon}\cdot(\dot{\boldsymbol{\varepsilon}}\wedge\boldsymbol{\tau}) + i\eta_4(\boldsymbol{\varepsilon}\cdot\dot{\boldsymbol{\varepsilon}})(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}) + i\eta_5\dot{a}(\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}), \tag{C.3}$$

with

$$\begin{aligned}\eta_1 &= (\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon})(\chi_1'' - 4g^2\chi_1 + 4g\chi_2 + 2(1+2f)\chi_3' + 2f'\chi_3 - 4g(\varepsilon\cdot n)\chi_4)\\ &\quad + 2(\varepsilon\cdot n)^2\chi_4 - (3\chi_1' + 2\chi_2 + 4(1+2f)\chi_3) - (3 + 8f(1+f))\chi_1,\\ \eta_2 &= (\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon})(\chi_2'' - 8g^2\chi_2 - 2\chi_4') + (\varepsilon\cdot n)(-5\chi_2' + 8g\chi_2 + 8\chi_4) - 2\chi_1' + 8g\chi_1\\ &\quad - 2(3 + 4f(1+f))\chi_2 + 4(1+2f)\chi_3,\\ \eta_3 &= (\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon})(\chi_3'' - 4g^2\chi_3) + (\varepsilon\cdot n)(-5\chi_3' + 4g\chi_3 - 2(1+2f)\chi_4)\\ &\quad - 2(1+2f)\chi_1' - 2f'\chi_1 + 2(1+2f)\chi_2 - 4(1 + 2f(1+f))\chi_3,\\ \eta_4 &= (\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon})(\chi_4'' - 8g^2\chi_4) + (\varepsilon\cdot n)(12g\chi_4 - 7\chi_4') - 4g^2\chi_1 + 2\chi_2' - 4g\chi_2\\ &\quad - 2(1+2f)\chi_3' - 2f'\chi_3 - (5 + 8f(1+f))\chi_4,\\ \eta_5 &= (\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon})(\chi_5'' - 8g^2\chi_5) + (\varepsilon\cdot n)(8g\chi_5 - 5\chi_5') - (3 + 8f(1+f))\chi_5,\end{aligned} \tag{C.4}$$

where χ_i depend on the S^3 coordinates through $(\varepsilon\cdot n)$ only (which holds equally well for f, g and h). The prime denotes derivation with respect to $(\varepsilon\cdot n)$. We can truncate η_i to the same order as was done for χ_i in eq. (C.2), and introduce the nine-dimensional basis:

$$Z_\mu = \{\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}, \varepsilon\cdot n\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}, (\varepsilon\cdot n)^2\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{\tau}, \dot{\varepsilon}\cdot n\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}, \varepsilon\cdot n\dot{\varepsilon}\cdot n\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}, \boldsymbol{\varepsilon}\cdot(\dot{\boldsymbol{\varepsilon}}\wedge\boldsymbol{\tau}), \varepsilon\cdot n\boldsymbol{\varepsilon}\cdot(\dot{\boldsymbol{\varepsilon}}\wedge\boldsymbol{\tau}), \boldsymbol{\varepsilon}\cdot\dot{\boldsymbol{\varepsilon}}\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}, \dot{a}\boldsymbol{\varepsilon}\cdot\boldsymbol{\tau}\}. \tag{C.5}$$

With respect to this basis $D_i\dot{A}_i = b_\mu Z_\mu$, with b_μ read off from eq. (C.2). For the moduli space

$$f_0 = -a/2, \quad g_0 = 1/2, \quad f_1 = a/2, \quad g_1 = -1/2. \tag{C.6}$$

Taking $\chi_1 = c_1 + c_2(\varepsilon\cdot n) + c_3(\varepsilon\cdot n)^2$, $\chi_2 = c_4 + c_5(\varepsilon\cdot n)$, $\chi_3 = c_6 + c_7(\varepsilon\cdot n)$, $\chi_4 = c_8$ and $\chi_5 = c_9$, eq. (C.3) can be written as

$$D_l^2(c_\mu Z_\mu) = Z_\mu \mathcal{A}_{\mu\nu} c_\nu, \tag{C.7}$$

After some tedious algebra one finds for the 9×9 matrix $\mathcal{A}$

$$\begin{pmatrix}
-(1+2a^2+\varepsilon^2) & 0 & 2\varepsilon^2 & -2\varepsilon^2 & 0 & a\varepsilon^2 & -2a\varepsilon^2 & 0 & 0 \\
4a^2 & -2(2+a^2) & 0 & -2 & 0 & 4a & 0 & 0 & 0 \\
1-2a^2+8af_0 & 4a^2 & -9-2a^2 & 2 & -2 & -5a & 6a & 2 & 0 \\
-4 & -2 & 0 & -2(2+a^2) & 0 & -4a & 0 & 0 & 0 \\
8g_0 & -4 & -4 & -4(1-a^2) & -9-2a^2 & 4a & -4a & 8 & 0 \\
-a & 2a & 0 & -2a & 0 & -2(1+a^2) & 0 & 0 & 0 \\
-4f_0 & -3a & 4a & 2a & -2a & -2+4a^2 & -7-2a^2 & 2a & 0 \\
-1 & 0 & 0 & 2 & 2 & -a & 2a & -3-2a^2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1-2a^2
\end{pmatrix} \tag{C.8}$$

which can be inverted, if necessary, with an algebraic manipulation programme. We will not present that result here. Using

$$\frac{1}{2\pi^2}\int_{S^3}\mathrm{Tr}(Z_\mu Z_\nu) = \begin{pmatrix}
L_1 & 0 & L_2 & 0 & L_3 & 0 & 0 & L_5 & L_6 \\
0 & L_2 & 0 & L_3 & 0 & 0 & 0 & 0 & 0 \\
L_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & L_3 & 0 & L_2 & 0 & 0 & 0 & 0 & 0 \\
L_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & L_4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & L_7
\end{pmatrix} \equiv L_{\mu\nu}, \tag{C.9}$$

with

$$\begin{aligned}
L_1 &= \tfrac{3}{2}\dot{\varepsilon}_\mu^2, \\
L_2 &= \tfrac{5}{12}\varepsilon_\mu^2\dot{\varepsilon}_\nu^2 - \tfrac{1}{6}(\varepsilon_\mu\dot{\varepsilon}_\mu)^2, \\
L_3 &= \tfrac{1}{3}\varepsilon_\mu^2\dot{\varepsilon}_\nu^2 - \tfrac{1}{12}(\varepsilon_\mu\dot{\varepsilon}_\mu)^2 \\
L_4 &= \varepsilon_\mu^2\dot{\varepsilon}_\nu^2 - (\varepsilon_\mu\dot{\varepsilon}_\mu)^2, \\
L_5 &= \tfrac{7}{6}(\varepsilon_\mu\dot{\varepsilon}_\mu)^2 + \tfrac{1}{12}\varepsilon_\mu^2\dot{\varepsilon}_\nu^2, \\
L_6 &= \tfrac{3}{2}\dot{a}\varepsilon_\mu\dot{\varepsilon}_\mu, \\
L_7 &= \tfrac{3}{2}\dot{a}^2\varepsilon_\mu^2,
\end{aligned} \tag{C.10}$$

(only those terms were kept that will contribute to $g_{\mu\nu}$, $g_{0\mu}$ and g_{00} to second order in ε) one straightforwardly computes

$$-\int_{S^3} \mathrm{Tr}(D_i \dot{A}_i D_l^{-2} D_j \dot{A}_j) = b_\mu L_{\mu\nu} A^{-1}_{\nu\lambda} b_\lambda . \tag{C.11}$$

The computation of $-\int_{S^3} \mathrm{Tr}(\dot{A}_i^2)$, which is left to the reader, completes the evaluation of the metric. The results are collected in eqs. (50) and (51).

We define the riemannian curvature by

$$R^\mu_{\nu\lambda\sigma} = \partial_\lambda \Gamma^\mu_{\nu\sigma} - \partial_\sigma \Gamma^\mu_{\lambda\nu} + \Gamma^\mu_{\lambda\alpha}\Gamma^\alpha_{\nu\sigma} - \Gamma^\mu_{\sigma\alpha}\Gamma^\alpha_{\nu\lambda} . \tag{C.12}$$

It is most easy to extract the Christoffel symbols from the equations of motion ($x^0 = a$, $x^\mu = \varepsilon^\mu$) derived from the metric tensor

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu \dot{x}^\lambda = 0 . \tag{C.13}$$

This fixes all our conventions. Since we will compute R for $\varepsilon = 0$, it is sufficient to evaluate Γ to first order in ε. As this computation is standard, it will not be reproduced here. The result is collected in eqs. (56)–(58).

Appendix D

This appendix collects the proofs of eqs. (85), (92) and (93). The following identities will be useful:

$$L_1^a n_\mu = \frac{i}{2} e^a_\mu = \frac{i}{2}\eta^a_{\mu\nu} n_\nu = \tfrac{1}{8}\{\sigma_\mu \overline{\sigma}\cdot n - \sigma\cdot n \overline{\sigma}_\mu, \tau_a\} , \tag{D.1}$$

$$L_2^a n_\mu = \frac{i}{2} \overline{e}^a_\mu = \frac{i}{2}\overline{\eta}^a_{\mu\nu} n_\nu = \tfrac{1}{8}\{\overline{\sigma}_\mu \sigma\cdot n - \overline{\sigma}\cdot n \sigma_\mu, \tau_a\} , \tag{D.2}$$

$$\sigma_\mu x \sigma_\mu = \overline{\sigma}_\mu x \overline{\sigma}_\mu = -2\overline{x} , \qquad x \equiv x_\lambda \sigma_\lambda , \tag{D.3}$$

$$\sigma_\mu x \overline{\sigma}_\mu = \overline{\sigma}_\mu x \sigma_\mu = 2(x + \overline{x}) = 4\mathrm{Re}\, x . \tag{D.4}$$

For eq. (D.1), see eqs. (4), (5) and (20), for eq. (D.2) also eqs. (12) and (21) are relevant. Eqs. (D.3) and (D.4) are proven by using explicitly the SU(2) algebra properties. Note that eq. (D.3) applies also to a product of quaternions, as this is again a quaternion. Furthermore, eq. (D.4) implies

$$\sigma_\mu \tau_a \overline{\sigma}_\mu = \sigma_\mu \sigma_a \overline{\sigma}_\mu = 0 . \tag{D.5}$$

We now turn to the proof of eq. (85):

$$\begin{aligned} L_1^a n\cdot\overline{\sigma}\tau_b n\cdot\sigma &= \frac{i}{2}(\overline{\sigma}_\mu e^a_\mu \tau_b n\cdot\sigma + \text{h.c.}) \\ &= \tfrac{1}{8}\overline{\sigma}_\mu\{\sigma_\mu\overline{\sigma}\cdot n - \sigma\cdot n\overline{\sigma}_\mu, \sigma_a\}\overline{\sigma}_b n\cdot\sigma - \text{h.c.} \\ &= \tfrac{1}{8}(6\overline{\sigma}\cdot n\sigma_a\overline{\sigma}_b n\cdot\sigma + 2(\sigma_a\sigma\cdot n)^\dagger\overline{\sigma}_b n\cdot\sigma) - \text{h.c.} \\ &= \tfrac{1}{2}\overline{\sigma}\cdot n[\tau_a,\tau_b]\sigma\cdot n = i\varepsilon_{abc}\overline{\sigma}\cdot n\tau_c\sigma\cdot n . \end{aligned} \tag{D.6}$$

Next we address the proof of eq. (93)

$$
\begin{aligned}
L_2^a n \cdot \overline{\sigma}\tau_b n \cdot \sigma &= \frac{i}{2}(\overline{e}_\mu^a \overline{\sigma}_\mu \tau_b n \cdot \sigma + \text{h.c.}) \\
&= \tfrac{1}{8}\{\overline{\sigma}_\mu \sigma \cdot n - \overline{\sigma} \cdot n \sigma_\mu, \sigma_a\} \overline{\sigma}_\mu \overline{\sigma}_b n \cdot \sigma - \text{h.c.} \\
&= \tfrac{1}{8}(-2(\sigma \cdot n \sigma_a)^\dagger - 6\sigma_a \overline{\sigma} \cdot n)\overline{\sigma}_b n \cdot \sigma - \text{h.c.} \\
&= -\tfrac{1}{2}[\tau_a, n \cdot \overline{\sigma}\tau_b n \cdot \sigma] . \qquad \text{(D.7)}
\end{aligned}
$$

Finally we prove the second identity in eq. (91). From the isotropy of S^3 one easily deduces

$$
\frac{1}{2\pi^2}\int_{S^3} n_\mu n_\nu = \tfrac{1}{4}\delta_{\mu\nu}, \quad \frac{1}{2\pi^2}\int_{S^3} n_\mu n_\nu n_\lambda n_\sigma = \tfrac{1}{24}(\delta_{\mu\nu}\delta_{\lambda\sigma} + \delta_{\mu\lambda}\delta_{\nu\sigma} + \delta_{\mu\sigma}\delta_{\nu\lambda}) , \qquad \text{(D.8)}
$$

such that

$$
\begin{aligned}
\frac{1}{2\pi^2}\int_{S^3} V_a^b V_c^d &= \tfrac{1}{24}\mathrm{Tr}(\overline{\sigma}_\mu \frac{\tau_a}{2}\sigma_\mu \tau_b)\mathrm{Tr}(\overline{\sigma}_\nu \frac{\tau_c}{2}\sigma_\nu \tau_d) \\
&+ \tfrac{1}{24}\mathrm{Tr}(\overline{\sigma}_\mu \frac{\tau_a}{2}\sigma_\nu \tau_b)\mathrm{Tr}(\overline{\sigma}_\nu \frac{\tau_c}{2}\sigma_\mu \tau_d) \\
&+ \tfrac{1}{24}(\overline{\sigma}_\mu \frac{\tau_a}{2}\sigma_\nu \tau_b)\mathrm{Tr}(\overline{\sigma}_\mu \frac{\tau_c}{2}\sigma_\nu \tau_d) . \qquad \text{(D.9)}
\end{aligned}
$$

The first term vanishes (see eq. (D.5)), on the second we can apply the completeness condition of eq. (84), whereas for the third we use that $\sigma_\nu \otimes \sigma_\nu = -\sigma_\nu \otimes \overline{\sigma}_\nu + 2\sigma_0 \otimes \sigma_0$. For $\sigma_\nu \otimes \overline{\sigma}_\nu$ we can again use eq. (84). Therefore,

$$
\begin{aligned}
\frac{1}{2\pi^2}\int_{S^3} V_a^b V_c^d &= \tfrac{1}{48}\mathrm{Tr}(\sigma_b \overline{\sigma}_\mu \sigma_a \sigma_c \sigma_\mu \sigma_d) \\
&- \tfrac{1}{48}\mathrm{Tr}(\sigma_b \overline{\sigma}_\mu \sigma_a \sigma_d \overline{\sigma}_\mu \sigma_c) \\
&+ \tfrac{1}{48}\mathrm{Tr}(\sigma_a \sigma_b \overline{\sigma}_\mu)\mathrm{Tr}(\sigma_c \sigma_d \overline{\sigma}_\mu) . \qquad \text{(D.10)}
\end{aligned}
$$

As $\overline{\sigma}_\mu \otimes \overline{\sigma}_\mu = \sigma_\mu \otimes \sigma_\mu$, we can repeat the last step on the last term. Using also eqs. (D.3) and (D.4) we find

$$
\begin{aligned}
\frac{1}{2\pi^2}\int_{S^3} V_a^b V_c^d &= \tfrac{1}{24}\mathrm{Tr}(\sigma_b(\sigma_a\sigma_c + (\sigma_a\sigma_c)^\dagger)\sigma_d) + \tfrac{1}{24}\mathrm{Tr}(\sigma_b\sigma_d\sigma_a\sigma_c) \\
&- \tfrac{1}{24}\mathrm{Tr}(\sigma_a\sigma_b\sigma_c\sigma_d) + \tfrac{1}{24}\mathrm{Tr}(\sigma_a\sigma_b)\mathrm{Tr}(\sigma_c\sigma_d) \\
&= \tfrac{1}{6}\delta_{ac}\delta_{bd} - \tfrac{1}{24}\mathrm{Tr}([\sigma_b, \sigma_c]\sigma_d\sigma_a) + \tfrac{1}{6}\delta_{ab}\delta_{cd} = \tfrac{1}{3}\delta_{ac}\delta_{bd} . \qquad \text{(D.11)}
\end{aligned}
$$

References

[1] G. 't Hooft, Phys. Rev. D14 (1976) 3434; [Erratum: D18 (1987) 2199]; Phys. Rev. Lett. 37 (1976) 8; Phys. Rep. 142 (1986) 357
[2] S. Coleman, *in*: The whys of subnuclear physics, ed. A. Zichichi (Plenum, New York, 1977)

[3] A. Ringwald, Nucl. Phys. B330 (1990) 1;
O. Espinosa, Nucl. Phys. B343 (1990) 310;
L. McLerran, A. Vainshtein and M. Voloshin, Phys. Rev. D42 (1990) 171
[4] F. Klinkhamer and N. Manton, Phys. Rev. D30 (1984) 2212
[5] C. Callan, R. Dashen and D. Gross, Phys. Rev. D17 (1978) 2717
[6] M. Lüscher, Nucl. Phys. B219 (1983) 233;
M. Lüscher and G. Münster, Nucl. Phys. B232 (1984) 445
[7] R. Cutkosky and K. Wang, Phys. Rev. D36 (1987) 3825; D37 (1988) 3024
[8] J. Koller and P. van Baal, Nucl. Phys. B302 (1988) 1;
P. van Baal, Acta Phys. Pol. B20 (1989) 295; Phys. Lett. B224 (1989) 397; Nucl. Phys. B351 (1991) 183.
[9] G. 't Hooft, Nucl. Phys. B153 (1979) 141; Acta Phys. Austr. Suppl. 22 (1980) 531
[10] C. Michael, G. Tickle and M. Teper, Phys. Lett. B207 (1988) 313;
C. Michael, Nucl. Phys. B339 (1990) 225
[11] C. Vohwinkel, Phys. Rev. Lett. 63 (1989) 2544; Ph. D. thesis, Tallahassee, September 1989
[12] P. van Baal, Nucl. Phys. B264 (1986) 548;
J. Koller and P. van Baal, Nucl. Phys. B273 (1986) 387; Ann. Phys. (NY) 174 (1987) 299
[13] T. Banks, G. Farrar, M. Dine, D. Karabali and B. Sakita, Nucl. Phys. B347 (1990) 581;
S. Khlebnikov, V. Rubakov and P. Tinyakov, Nucl. Phys. B347 (1990) 783;
J. Cornwall, Phys. Lett. B243 (1990) 271;
V. Zakharov, Munich preprints MPI-PAE/PTh 52/90 and MPI-PAE/PTh 11/91
[14] J. Hoek, Nucl. Phys. B332 (1990) 530
[15] H. Schenk, Commun. Math. Phys. 116 (1988) 177;
P. Braam and P. van Baal, Commun. Math. Phys. 122 (1989) 267
[16] G. Savvidy, Phys. Lett. B71 (1977) 133;
N. Nielsen and P. Oleson, Nucl. Phys. B144 (1978) 376;
T. Hansson, K. Johnson and C. Peterson, Phys. Rev. D26 (1982) 2069;
T. Hansson, P. van Baal and I. Zahed, Nucl. Phys. B289 (1987) 682
[17] J. Koller and P. van Baal, Nucl. Phys. B (Proc. Suppl.) 4 (1988) 47
[18] P. van Baal, *in* Probalilistic methods in quantum field theory and quantum gravity, ed. P. Damgaard et al. (Plenum, New York, 1990) p. 131
[19] P. Nelson and L. Alvarez-Gaumé, Commun. Math. Phys. 99 (1985) 103
[20] P. van Baal, Nucl. Phys. B369 (1992) 259
[21] A. Belavin, A. Polyakov, A. Schwarz and Y. Tyupkin, Phys. Lett. B59 (1975) 85;
M. Atiyah, V. Drinfeld, N. Hitchin and Yu. Manin, Phys. Lett. A65 (1978) 185
[22] N. Manton, Phys. Lett. B110 (1982) 54; B154 (1985) 397;
B. Schroer, Nucl. Phys. B367 (1991) 177
[23] M. Atiyah and N. Hitchin, Phys. Lett. A107 (1985) 21; The geometry and dynamics of magnetic monopoles (Princeton University Press, 1989)
[24] I. Balitsky and A. Yung, Phys. Lett. B168 (1986) 113;
A. Yung, Nucl. Phys. B297 (1988) 47;
E. Shuryak, Nucl. Phys. B302 (1988) 559;
J. Verbaarschot, Nucl. Phys. B362 (1991) 33;
V. Khoze and A. Ringwald, Cern preprint, Cern-TH.6082/91, May 1991
[25] C. Bloch, Nucl. Phys. 6 (1958) 329
[26] M. Atiyah, The geometry of Yang-Mills fields, Fermi lectures (Scuola Normale, Pisa, 1979)
[27] P. van Baal, Tunnelling and the path decomposition expansion, lectures delivered at the Adriatico course on path integration, Utrecht preprint THU-91/19, October 1991
[28] O. Babelon and C. Viallet, Commun. Math. Phys. 81 (1981) 515
[29] S. Donaldson and P. Kronheimer, The geometry of four-manifolds (Oxford Univ. Press, Oxford, 1990);
D. Freed and K. Uhlenbeck, Instantons and four-manifolds, M.S.R.I. publications, Vol 1 (Springer, New York, 1984)
[30] M. Lüscher, Nucl. Phys. B205 [FS5] (1982) 483

[31] D. Groisser and T. Parker, Commun. Math. Phys. 112 (1987) 663; 135 (1990) 101
[32] H. Doi, Y. Matsumoto and T. Matumoto, *in* A fête of topology, ed. Y. Matsumoto et al, (Academic Press, New York, 1988) p. 543
[33] D. Groisser and T. Parker, J. Diff. Geom. 29 (1989) 499
[34] N. Manton, Phys. Rev. Lett. 60 (1988) 1916
[35] N. Christ and T.D. Lee, Phys. Rev. D22 (1980) 939
[36] G. Savvidy, Phys. Lett. B159 (1985) 325
[37] M. Semenov-Tyan-Shaskii and V. Franke, Proc. Steklov seminars 120 (1982) 159, English translation (Plenum, New York, 1986) p. 1999;
G. Dell'Antonio and D. Zwanziger, Commun. Math. Phys. 138 (1991) 107
[38] G. Dell'Antonio and D. Zwanziger, Nucl. Phys. B326 (1989) 333
[39] C. Vohwinkel, work in progress, private communication
[40] R. Cutkosky, Czech. J. Phys. 40 (1990) 252
[41] D. Zwanziger, Critical limit of lattice gauge theory, NYU preprint, December 1991
[42] S. Wolfram et al, Mathematica (Addison-Wesley, New York, 1991)
[43] P. van Baal and R.E. Cutkosky, Non-perturbative analysis, Gribov horizons and the boundary of the fundamental domain, Carnegie Mellon/Utrecht preprint CMU-HEP 92-19/TH-92/20 (August, 1992), Proc. XXI DGM (World Scientific, Singapore), to be published