Nuclear Physics B400 (1993) 126–144
North-Holland

NUCLEAR
PHYSICS B

# Yang–Mills theory without Mandelstam constraints

R. Loll [1]

*Physics Department, Syracuse University, Syracuse, NY 13244, USA*

Received 14 July 1992
(Revised 13 October 1992)
Accepted for publication 22 February 1993

We continue our investigation of the SU(2) lattice gauge theory in terms of an independent set of gauge-invariant Wilson loops. It is demonstrated how the standard measure of the gauge-covariant formulation can be transformed to a loop-variable measure. Explicit calculations in two and three dimensions illustrate the validity of our approach. We solve the $1 \times 1$ theory in terms of loop variables, and discuss the derivation of the classical equations of motion and possible consequences for the continuum theory.

## 1. Introduction

Path-dependent formulations of Yang–Mills theory, although often appealing heuristically, have not been very successful in solving outstanding problems of the theory. Attempts to describe gauge theory solely in terms of path-dependent field variables typically involve the holonomy along an open or closed path $\gamma$ in the space-time or spatial manifold $\Sigma^d$,

$$U_A(\gamma) := \mathrm{P} \exp \int_\gamma A_\mu \mathrm{d}\gamma^\mu, \tag{1.1}$$

with P denoting path ordering, $A_\mu$ the Yang–Mills gauge potential of a classical, semisimple Lie group G, and the dependence on the endpoints of $\gamma$ being understood. If $\gamma$ is a closed curve (a "loop"), one can define another non-local quantity, the traced holonomy

$$T_A(\gamma) := \mathrm{Tr}\, U_A(\gamma), \tag{1.2}$$

which is invariant under gauge transformations of $A$ and does not depend on which point of $\gamma$ is chosen as its basepoint. Unfortunately, such approaches tend to run into related problems which have their origin in the fact that the equations of the theory are not defined on $\Sigma^d$ itself, but on some infinite-dimensional space of paths or loops in $\Sigma^d$, which is mathematically ill-defined. Part of the problem is that, roughly speaking, there are many more paths in $\Sigma^d$ than there are loops,

[1] Supported by Deutsche Forschungsgemeinschaft

and as a consequence sets of path-dependent field variables contain a lot of redundant information.

It even seems problematic to write down meaningful classical equations of motion in terms of path-dependent variables; existing examples are flawed in various ways. For instance, in the non-local formulations of Bialynicki-Birula [1] and Mandelstam [2], Polyakov [3,4], and Migdal and Makeenko [5,6], the classical equations "equivalent" to the usual local Yang–Mills equations are all *linear* in the basic field variables. Obviously the non-trivialities of non-abelian gauge theories associated with their non-linear structure are hidden in these approaches; they reappear as problems to do with the structure of the space of paths or loops. Also note there is up to now no action principle for such non-local formulations, and the equations of motion therefore have to be derived by other, less direct means.

Our interest is the most "radical" kind of reformulation of gauge theory, one that dispenses with all gauge-*co*variant quantities, and uses exclusively traced holonomies as its fundamental variables. The physical motivation is the assumption that Yang–Mills theory "in the confined phase" can be described purely in terms of gauge-*in*variant variables. The mathematical justification for this approach is the proof of equivalence between the set $\mathcal{A}$ of all gauge potentials modulo the group $\mathcal{G}$ of local gauge transformations and the set of all traced holonomies $\{T(\gamma)\}$ (where $\gamma$ runs through all loops in $\Sigma^d$) modulo the Mandelstam constraints [7]. We write

$$\mathcal{A}/\mathcal{G} \simeq \{T(\gamma),\ \gamma \in \mathcal{L}\Sigma^d\}/\{\text{Mandelstam constraints}\}, \tag{1.3}$$

where $\mathcal{L}\Sigma^d$ denotes an appropriate quotient (with respect to orientation preserving reparametrizations and retracing equivalence) of the loop space of $\Sigma^d$ (for more details on loop spaces, see ref. [8]).

The existence of the Mandelstam constraints, an infinite set of G-dependent non-linear algebraic equations in terms of the traced holonomies $T(\gamma)$, is an expression of the overcompleteness of loop variables mentioned earlier. For example, for gauge group SU(2), we have a constraint

$$T_A(\alpha)T_A(\beta) - T_A(\alpha \circ_x \beta) - T_A(\alpha \circ_x \beta^{-1}) = 0 \tag{1.4}$$

for each configuration of two loops $\alpha$ and $\beta$ intersecting in a common point $x$ (with $\circ_x$ denoting the standard loop composition in $x$). Up to now nobody has found a way of implementing the full set of Mandelstam constraints in the classical theory. This means that loop-dependent formulations, although gauge-invariant, are *not* given in terms of the true physical degrees of freedom, but are again theories of constrained type. We do not know which sector of the large loop space is physically relevant and, consequently, which mathematical structures should best be imposed on it.

An area where loop space methods have been implemented to describe gauge field dynamics not just at a formal level, is the lattice gauge theory, where loop space is approximated by the discrete set of closed contours on a hypercubic finite space(-time) lattice. A number of authors have dealt with such formulations, mainly in the hamiltonian context (see for example refs. [9–13]). The number of physical degrees of freedom becomes finite, however, one is again faced with the issue of overcompleteness due to the presence of Mandelstam constraints, which impose non-linear identities between traced holonomies of lattice loops that intersect or share links in common. If those are only partially eliminated, the number of loop variables grows very rapidly with growing lattice size, thus obstructing computations for bigger lattices.

It was shown in ref. [14] that for the case of $G=SU(2)$ and in dimensions 2, 3 and 4, one can solve the Mandelstam constraints explicitly and thereby arrive at a local description in terms of an independent set of traced lattice holonomies, the so-called $L$-variables. The interesting features of this new description are its homogeneity throughout the lattice and the fact that one only needs to consider small loops on the lattice, extending over at most two lattice plaquettes.

Given this set of gauge-invariant, independent basic variables, a crucial question is whether the dynamics can be expressed in a manageable form in terms of them. This is where similar attempts of identifying the true degrees of freedom of Yang–Mills theory have typically run into trouble [16,17]. The present paper aims at demonstrating that it is actually possible to "do physics" in terms of the independent loop variables proposed in ref. [14]. Although no immediate obstructions appear in our formulation, we cannot exclude that difficulties may arise in the application to bigger lattices. We do however regard the approach as valuable in itself, since it introduces methods peculiar to a pure loop formulation, which differ in principle from those of the usual connection formulation, and may be of relevance for the corresponding continuum loop theory. We will be concerned here with the case $G=SU(2)$, but the corresponding treatment for other gauge groups, say, $SU(N)$, is expected to be in many regards similar.

In sect. 1, the discussion of the independent loop variables of ref. [14] is summarized and extended by the new relation (III), underlining the importance of a set of inequalities in the description of the reduced, physical configuration space, and commenting on some non-trivial topological properties of the description. Sect. 3 contains the explicit form of the holonomy matrices reconstructed from the independent $L$-variables for a set of loops intersecting at a common point. This is necessary for computing the measure over the physical configuration space in terms of the $L$-variables. This measure is computed locally for the case of the three-dimensional lattice. In sect. 4, these results are used to derive the explicit partition function and the one-plaquette expectation value for the $1\times 1$-lattice, demonstrating the feasibility of the approach. The non-trivial topological features of this example are analyzed completely. Sect. 5 consists of an essentially independent discussion of how the classical lattice loop equations of motion can be derived consistently from an action principle, although the Wilson lagrangian is linear in the basic variables. To this end it is convenient to re-express the $L$-variables in terms of a set of angular variables, exhibiting the non-trivial global structure of the physical configuration space. These results are relevant for a pure loop approach in the continuum and for a hamiltonian version of the formulation in terms of independent lattice loops. Sect. 6 contains the conclusions and discusses possible consequences for an analogous loop description of the continuum theory.

## 2. Independent loop variables revisited

As has been illustrated in ref. [14], a particularly convenient set of loop variables for the gauge-invariant description of SU(2) Yang–Mills theory is given in terms of the so-called $L$-variables, which are certain linear combinations of the traced holonomies (1.2). Consider a set of oriented loops $\alpha, \beta, \gamma, \ldots$ in the $d$-dimensional manifold $\Sigma^d$, all starting and ending at a fixed basepoint $x_0 \in \Sigma^d$. We define three types of loop variables $L_i$ with $i$ loop arguments:

$$\begin{aligned}
L_1(\alpha) &:= \tfrac{1}{2} T(\alpha)\,, \\
L_2(\alpha, \beta) &:= \tfrac{1}{4}\Big(T(\alpha\circ\beta^{-1}) - T(\alpha\circ\beta)\Big)\,, \\
L_3(\alpha,\beta,\gamma) &:= \tfrac{1}{16}\Big(T(\alpha\circ\gamma\circ\beta) - T(\alpha\circ\gamma^{-1}\circ\beta) \\
&\quad - T(\alpha\circ\gamma\circ\beta^{-1}) + T(\alpha\circ\gamma^{-1}\circ\beta^{-1}) \\
&\quad - T(\alpha\circ\beta\circ\gamma) + T(\alpha\circ\beta^{-1}\circ\gamma) \\
&\quad + T(\alpha\circ\beta\circ\gamma^{-1}) - T(\alpha\circ\beta^{-1}\circ\gamma^{-1})\Big)\,,
\end{aligned} \tag{2.1}$$

where it is understood that the loops appearing in the arguments of the traced holonomies $T$ are composed at $x_0$. We have slightly changed notation and normalization of the $L$-variables, compared with the definition given in ref. [14]. Using the explicit parametrization for the SU(2) holonomy matrix in the fundamental representation, for a loop $\alpha$,

$$U_\alpha = \begin{pmatrix} \alpha_1 + i\alpha_2 & \alpha_3 + i\alpha_4 \\ -\alpha_3 + i\alpha_4 & \alpha_1 - i\alpha_2 \end{pmatrix}, \tag{2.2}$$

by four real parameters $\alpha_i$, subject to the constraint $\sum_i \alpha_i^2 = 1$, we obtain

$$\begin{aligned}
L_1(\alpha) &= \alpha_1\,, \\
L_2(\alpha,\beta) &= \alpha_2\beta_2 + \alpha_3\beta_3 + \alpha_4\beta_4 = \boldsymbol{\alpha}_\perp \cdot \boldsymbol{\beta}_\perp\,, \\
L_3(\alpha,\beta,\gamma) &= \boldsymbol{\gamma}_\perp \cdot (\boldsymbol{\alpha}_\perp \times \boldsymbol{\beta}_\perp)\,.
\end{aligned} \tag{2.3}$$

The vectorial notation $\boldsymbol{\alpha}_\perp$ stands for the 2-, 3- and 4-component of the unit vector $\boldsymbol{\alpha}$ in $\mathbb{R}^4$. As already mentioned in the previous section, a central problem of formulating Yang–Mills theory in terms of the gauge-invariant traced holonomies is the vast overcompleteness of these loop variables, due to the presence of the Mandelstam constraints. The $L$-variables of (2.1) get rid of some of this overcompleteness. It was demonstrated in ref. [14] how the traced holonomy $T(\alpha_1 \circ \alpha_2 \circ \alpha_3 \circ \dots)$ of an arbitrary product of oriented loops based at $x_0$, $\alpha_1 \circ \alpha_2 \circ \alpha_3 \circ \dots$, can be expressed as a function of the variables (2.3).

However, if one wants to re-express the theory in terms of these $L$-variables, one has to take into account they are still not independent, but obey certain polynomial constraints quadratic in $L$ [14], which again derive directly from the Mandelstam constraints (1.4). The one relevant to the present discussion is

$$L_2(\alpha,\alpha)L_2(\beta,\beta)L_2(\gamma,\gamma) - L_2(\alpha,\alpha)L_2(\beta,\gamma)^2 - L_2(\beta,\beta)L_2(\alpha,\gamma)^2 \\ -L_2(\gamma,\gamma)L_2(\alpha,\beta)^2 + 2L_2(\alpha,\beta)L_2(\alpha,\gamma)L_2(\beta,\gamma) = L_3(\alpha,\beta,\gamma)^2, \tag{2.4}$$

where the $L_2$ variable of two identical arguments is a function of $L_1$,

$$L_2(\alpha,\alpha) = 1 - L_1(\alpha)^2. \tag{2.5}$$

Eq. (2.5) implies the $L_3$ variables are determined by the $L_1$ and $L_2$ variables up to a sign.

Furthermore, there are inequalities that do not affect the number of local degrees of freedom, but restrict the allowed range of the $L$-variables to certain closed subsets of the real line. They follow immediately from the definition of the $L$-variables. We have

$$-1 \leqslant L_i \leqslant 1, \qquad i = 1,2,3, \tag{I}$$

$$-\sqrt{L_2(\alpha,\alpha)L_2(\beta,\beta)} \leqslant L_2(\alpha,\beta) \leqslant \sqrt{L_2(\alpha,\alpha)L_2(\beta,\beta)}. \tag{II}$$

Another inequality follows from eq. (2.4), namely,

$$L_2(\alpha,\alpha)L_2(\beta,\beta)L_2(\gamma,\gamma) - L_2(\alpha,\alpha)L_2(\beta,\gamma)^2 - L_2(\beta,\beta)L_2(\alpha,\gamma)^2 \\ -L_2(\gamma,\gamma)L_2(\alpha,\beta)^2 + 2L_2(\alpha,\beta)L_2(\alpha,\gamma)L_2(\beta,\gamma) \geqslant 0, \tag{III}$$

which can be rewritten as a condition of positive semi-definiteness of a determinant,

$$\det L := \begin{vmatrix} L_2(\alpha,\alpha) & L_2(\alpha,\beta) & L_2(\alpha,\gamma) \\ L_2(\alpha,\beta) & L_2(\beta,\beta) & L_2(\beta,\gamma) \\ L_2(\alpha,\gamma) & L_2(\beta,\gamma) & L_2(\gamma,\gamma) \end{vmatrix} \geqslant 0. \tag{III'}$$

This important inequality has not appeared in the literature so far. Note that (II) follows from (III) unless $\det L = 0$.

In the continuum theory, no explicit solution to the Mandelstam constraints is known, i.e. we do not know of a set of algebraically independent traced holonomies which could be used to parametrize the reduced configuration space

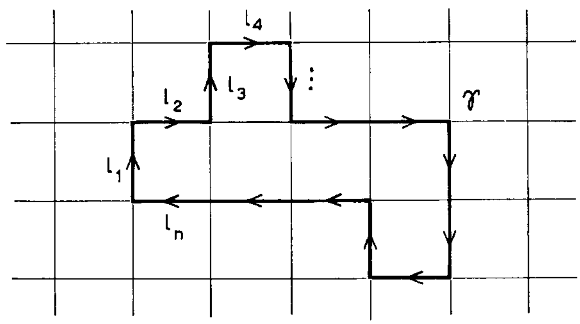



Fig. 1.

$\mathcal{A}/\mathcal{G}$. Even if such a description could be found locally, we do not expect such coordinates to exist *globally*, due to the non-trivial topological structure of $\mathcal{A}/\mathcal{G}$.

However, one can make progress toward solving this problem by looking at a regularized version of the theory, with the continuum structure of space-time approximated by a hypercubic lattice with periodic boundary conditions. The gauge-invariant observables in this case are the traced holonomies of lattice loops, i.e. closed contours consisting of chains of contiguous oriented links $l$ on the lattice (fig. 1). It was shown in ref. [14] that in $d = 2, 3, 4$ there exists a finite subset $\{\gamma\}$ of these lattice loops such that their associated traced holonomies $\{T(\gamma)\}$ give a good local parametrization of the reduced, *physical* configuration space

$$Q_{\text{red}} = \frac{\times_l \text{SU}(2)_l}{\times_s \text{SU}(2)_s} \tag{2.6}$$

of the standard lattice formulation (see for example ref. [18]), with $\times_l$ and $\times_s$ denoting the product over all links and sites respectively. The solution is surprisingly simple, although its derivation is not (it amounts to solving the highly coupled set of non-linear Mandelstam constraints on the lattice): the set $\{\gamma\}$ in all cases can be chosen to consist of short loops, going around one or at most two lattice plaquettes, corresponding to a description in terms of a subset of the $L_1$ and $L_2$ variables only. Moreover, in contrast with covariant, maximally gauge-fixed formulations, the choice of these loop variables is invariant under lattice translations.

Note that discrete ($\mathbb{Z}_2$) degrees of freedom get lost when transforming from the link holonomies modulo gauge transformations to the $L$-variables, because the latter depend on "contractible" loops only (loops with vanishing winding number on the discretized torus $\text{T}^d$), whereas certain gauge-inequivalent configurations of link holonomies can only be distinguished by considering traced holonomies of lattice loops with non-vanishing winding number, so-called

Polyakov loops *. The relation between the $L$-variables and these Polyakov loops will be illustrated with the help of the $1 \times 1$-lattice in sect. 4.

Not much is known about the topology of the reduced configuration space (2.6) of the lattice theory, for instance, its orbit structure, and not all of it will be relevant, since we are ultimately interested in the (quantum) continuum limit of the theory. The main contributions in this limit come from configurations with small action, which are close to the identity, $U_l \equiv 1\!\!1$. For $d = 4$, Lüscher has shown that, although every configuration of link holonomies can be deformed in a continuous way to the identity, a non-trivial topological charge can be recovered by restricting oneself to such small configurations [19]. Also note there are no unique lattice analogues of topological charges defined in the continuum theory on the torus $T^d$.

The existence of the independent loop variables is by itself a remarkable fact, but it is of crucial importance whether physically interesting quantities, such as the lagrangian and hamiltonian or the measures on configuration and phase space assume a manageable form when expressed in terms of them. First results, presented below, indicate this may indeed be true (to the extent one expects a formulation of non-abelian gauge theory to be "easy"). In this new formulation, the lagrangian is a trivial linear function in the basic variables, and the non-trivialities associated with the non-linearities of Yang–Mills theory show up in a different, and to our mind particularly transparent way, because throughout one works directly on the physical configuration space, without having to care about gauge covariance and/or the overcompleteness of loop variables. To what extent the lattice results may also change our view of the continuum theory will be discussed in sect. 6.

## 3. Loop variable measure

Let us again consider a configuration of $n$ non-selfintersecting ("basic") continuum loops $\gamma_i$, $i = 1 \dots n$, based and composed at the point $x_0 \in \Sigma$. Following Giles [7], we will give the explicit form of a set of $n$ SU(2)-matrices $U_{\gamma_i}$ in the fundamental representation, such that their traces and the traces of products of matrices $U$ are given in terms of the $L$-variables, i.e. a solution to the inverse problem of reconstructing the holonomies from the knowledge of a restricted subset of the traced holonomies.

Combining the results of refs. [7] and [14], we will arrive at explicit matrix expressions for the holonomies which again highlight the significance of the $L$-variables, in particular the role played by the three-loop variables $L_3$. Giles' construction involves selecting two reference loops $\alpha$ and $\beta$ from the set of the $\gamma_i$ and to represent $U_\alpha$ by a diagonal matrix and $U_\beta$ by a matrix with a particular

* I thank R. Ben-Av for pointing this out.

non-vanishing off-diagonal element. The freedom involved in choosing these two matrices for the case of SU(2) is reduced to

$$U_\alpha = \begin{pmatrix} L_1(\alpha) + i\epsilon\sqrt{L_2(\alpha,\alpha)} & 0 \\ 0 & L_1(\alpha) - i\epsilon\sqrt{L_2(\alpha,\alpha)} \end{pmatrix} \tag{3.1}$$

and

$$U_\beta = \begin{pmatrix} L_1(\beta) + i\epsilon\dfrac{L_2(\alpha,\beta)}{\sqrt{L_2(\alpha,\alpha)}} & \mathrm{e}^{i\phi}\sqrt{L_2(\beta,\beta) - \dfrac{L(\alpha,\beta)^2}{L_2(\alpha,\alpha)}} \\ -\mathrm{e}^{-i\phi}\sqrt{L_2(\beta,\beta) - \dfrac{L(\alpha,\beta)^2}{L_2(\alpha,\alpha)}} & L_1(\beta) - i\epsilon\dfrac{L_2(\alpha,\beta)}{\sqrt{L_2(\alpha,\alpha)}} \end{pmatrix}, \tag{3.2}$$

with arbitrary $\phi \in [0, 2\pi]$ and $\epsilon = \pm 1$. The matrix $U_{\gamma_i}$ representing any of the other loops $\gamma_i$ can now be constructed from the knowledge of $L_1(\gamma_i)$, $L_2(\alpha,\gamma_i)$, $L_2(\beta,\gamma_i)$ and $L_3(\alpha,\beta,\gamma_i)$ according to

$$U_{\gamma_i} = \begin{pmatrix} L_1(\gamma_i) + i\epsilon\dfrac{L_2(\alpha,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}} & \mathrm{e}^{i\phi}\dfrac{\left(L_2(\beta,\gamma_i) - \frac{L_2(\alpha,\beta)L_2(\alpha,\gamma_i)}{L_2(\alpha,\alpha)} + i\epsilon\frac{L_3(\alpha,\beta,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}}\right)}{\sqrt{L_2(\beta,\beta) - \frac{L_2(\alpha,\beta)^2}{L_2(\alpha,\alpha)}}} \\ \mathrm{e}^{-i\phi}\dfrac{\left(-L_2(\beta,\gamma_i) + \frac{L_2(\alpha,\beta)L_2(\alpha,\gamma_i)}{L_2(\alpha,\alpha)} + i\epsilon\frac{L_3(\alpha,\beta,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}}\right)}{\sqrt{L_2(\beta,\beta) - \frac{L_2(\alpha,\beta)^2}{L_2(\alpha,\alpha)}}} & L_1(\gamma_i) - i\epsilon\dfrac{L_2(\alpha,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}} \end{pmatrix} \tag{3.3}$$

The degenerate case where we cannot find a pair $(\alpha,\beta)$ such that $U_\alpha$ is different from the unit matrix and $U_\beta$ has a non-zero off-diagonal matrix element $(U_\beta)_{12}$ is discussed in ref. [7], and presents no obstruction in principle to constructing the holonomy matrices. The arbitrariness of the parameter $\phi$ can be traced back to the freedom of transforming all matrices $U$ by a diagonal similarity transformation

$$U \mapsto \Omega U \Omega^{-1}, \qquad \Omega = \begin{pmatrix} \mathrm{e}^{i\omega} & \\ & \mathrm{e}^{-i\omega} \end{pmatrix},\ \omega \in \mathbb{R}, \tag{3.4}$$

which leaves all traced holonomies unchanged. One can eliminate the remaining gauge freedom by fixing the values of $\phi$ and $\epsilon$, and we will in the following set $\phi = 0$ and $\epsilon = 1$. However, one has to remember they are unphysical quantities and cannot be determined from the traced holonomies. We observed in sect. 2 that the knowledge of the $L_1$ and $L_2$ variables is "almost enough" to reconstruct all the traced holonomies, and therefore the holonomies themselves. Expression (3.3) illustrates how the ensuing sign ambiguity of $L_3$ variables manifests itself in terms of the holonomy matrices $U_{\gamma_i}$, namely, as a sign ambiguity of the 4-component of the corresponding unit 4-vector $\gamma_i$.

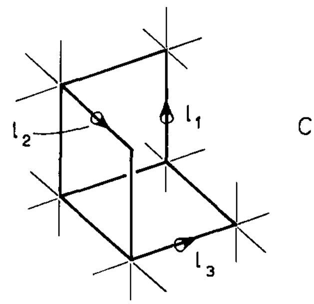



Fig. 2.

Note that for configurations with $L_3(\alpha,\beta,\gamma_i) = 0$ and with the above gauge choice for $\phi$ and $\epsilon$, (3.3) becomes

$$U_{\gamma_i} = \begin{pmatrix} L_1(\gamma_i) + i\dfrac{L_2(\alpha,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}} & \pm\sqrt{L_2(\gamma_i,\gamma_i) - \dfrac{L_2(\alpha,\gamma_i)^2}{L_2(\alpha,\alpha)}} \\ \mp\sqrt{L_2(\gamma_i,\gamma_i) - \dfrac{L_2(\alpha,\gamma_i)^2}{L_2(\alpha,\alpha)}} & L_1(\gamma_i) - i\dfrac{L_2(\alpha,\gamma_i)}{\sqrt{L_2(\alpha,\alpha)}} \end{pmatrix}. \quad (3.5)$$

Establishing explicit expressions for the holonomy matrices in terms of the $L$-variables is an important step in finding the correct measure for the lattice gauge theory in terms of independent loop variables. We illustrate in the following how this can be achieved.

Consider the case of the $N^3$ hypercubic lattice with periodic boundary conditions, with $N \geqslant 2$. Pick a particular $1\times1\times1$-unit cell $C$ on this lattice (in this example, the case $d = 3$ is easier than $d = 2$, since in three dimensions the loop variables are more localized). According to ref. [14], the six independent loop variables associated with $C$ can be expressed as functions of the holonomy variables of the nine links appearing in fig. 2. We choose a gauge-fixing for links such that only the three variables on the oriented links $l_1$, $l_2$ and $l_3$ remain unfixed (fig. 2). The corresponding oriented plaquette loops located on the faces of the unit cell $C$ we denote by $\alpha$, $\beta$ and $\gamma$. The six independent loop variables associated with this configuration are

$$\begin{aligned} L_1(\alpha) &= \tfrac{1}{2}\mathrm{Tr}\, U_\alpha = \tfrac{1}{2}\mathrm{Tr}\, U_{l_1}, \quad L_2(\alpha,\beta) = \tfrac{1}{4}(\mathrm{Tr}\, U_\alpha U_\beta^{-1} - \mathrm{Tr}\, U_\alpha U_\beta), \\ L_1(\beta) &= \tfrac{1}{2}\mathrm{Tr}\, U_\beta = \tfrac{1}{2}\mathrm{Tr}\, U_{l_2}, \quad L_2(\alpha,\gamma) = \tfrac{1}{4}(\mathrm{Tr}\, U_\alpha U_\gamma^{-1} - \mathrm{Tr}\, U_\alpha U_\gamma), \\ L_1(\gamma) &= \tfrac{1}{2}\mathrm{Tr}\, U_\gamma = \tfrac{1}{2}\mathrm{Tr}\, U_{l_3}, \quad L_2(\beta,\gamma) = \tfrac{1}{4}(\mathrm{Tr}\, U_\beta U_\gamma^{-1} - \mathrm{Tr}\, U_\beta U_\gamma). \end{aligned} \quad (3.6)$$

The standard choice for the measure in the integration over configuration space in terms of the link variables is just the product over all unfixed links $l_i$ of the invariant group measure for SU(2),

$$\int \mathrm{d}M = \prod_{l_i} \int \mathrm{d}M_{l_i}$$
$$= \prod_{l_i} \left( 2 \int \mathrm{d}\alpha_1 \int \mathrm{d}\alpha_2 \int \mathrm{d}\alpha_3 \int \mathrm{d}\alpha_4 \, \delta(\sum_{k=1}^{4} \alpha_k^2 - 1) \right)_{l_i}, \tag{3.7}$$

where the $\alpha_k$ are the coordinates of the realization of SU(2) as the unit sphere $S^3 \subset \mathbb{R}^4$, (2.2). In case we are integrating with this measure over a function which does not depend on the sign of $\alpha_4$, say, for a particular link $l$, the measure factor corresponding to $l$ reduces to

$$\int \mathrm{d}M_l = 2 \int_{-1}^{1} \mathrm{d}\alpha_1 \int_{-\sqrt{1-\alpha_1^2}}^{\sqrt{1-\alpha_1^2}} \mathrm{d}\alpha_2 \int_{-\sqrt{1-\alpha_1^2-\alpha_2^2}}^{\sqrt{1-\alpha_1^2-\alpha_2^2}} \mathrm{d}\alpha_3 \, \frac{1}{\sqrt{1-\alpha_1^2-\alpha_2^2-\alpha_3^2}}. \tag{3.8}$$

We will use the explicit forms (3.1), (3.2) and (3.3) to represent the matrices $U_{l_1} \equiv U_\alpha$, $U_{l_2} \equiv U_\beta$ and $U_{l_3} \equiv U_\gamma$ respectively, with gauge choice $\phi = 0$. Note that the sign of $L_3(\alpha, \beta, \gamma)$ in this case of just three basic loops is gauge because it is related to the sign of $\epsilon$.

We now perform a partial coordinate transformation from the canonical group coordinates associated with the links $l_1$, $l_2$ and $l_3$ to the $L$-variables (3.6). We can explicitly calculate the jacobian for this transformation. Although the special form of the three holonomy matrices is due to a particular gauge-fixing, the end result, when expressed in terms of the $L$-variables, is of course independent of this choice. One finds

$$\begin{aligned}
&\int \mathrm{d}M \\
&= 16\pi^2 \int \mathrm{d}L_1(\alpha) \int \mathrm{d}L_1(\beta) \int \mathrm{d}L_1(\gamma) \\
&\quad \times \int \mathrm{d}L_2(\alpha,\beta) \int \mathrm{d}L_2(\beta,\gamma) \int \mathrm{d}L_2(\alpha,\gamma) \int \mathrm{d}L_3(\alpha,\beta,\gamma) \\
&\quad \times \delta \Big( L_3(\alpha,\beta,\gamma)^2 - L_2(\alpha,\alpha) L_2(\beta,\beta) L_2(\gamma,\gamma) \\
&\qquad + L_2(\alpha,\alpha) L_2(\beta,\gamma)^2 + L_2(\beta,\beta) L_2(\alpha,\gamma)^2 \\
&\qquad + L_2(\gamma,\gamma) L_2(\alpha,\beta)^2 - 2 L_2(\alpha,\beta) L_2(\alpha,\gamma) L_2(\beta,\gamma) \Big) \\
&\quad \times \left( \prod_{l_i, i \neq 1,2,3} \int \mathrm{d}M_{l_i} \right)
\end{aligned}$$

$$= 16\pi^2 \int_{-1}^{1} \mathrm{d}L_1(\alpha) \int_{-1}^{1} \mathrm{d}L_1(\beta) \int_{-1}^{1} \mathrm{d}L_1(\gamma)$$

$$\times \int_{-\sqrt{L_2(\alpha,\alpha)L_2(\beta,\beta)}}^{\sqrt{L_2(\alpha,\alpha)L_2(\beta,\beta)}} \mathrm{d}L_2(\alpha,\beta) \int_{-\sqrt{L_2(\beta,\beta)L_2(\gamma,\gamma)}}^{\sqrt{L_2(\beta,\beta)L_2(\gamma,\gamma)}} \mathrm{d}L_2(\beta,\gamma) \int_{L_+}^{L_-} \mathrm{d}L_2(\alpha,\gamma)$$

$$\times \frac{1}{\sqrt{L_3(\alpha,\beta,\gamma)^2}} \left( \prod_{l_i, i\neq 1,2,3} \int \mathrm{d}M_{l_i} \right), \tag{3.9}$$

with

$$L_\pm = \frac{L_2(\alpha,\beta)L_2(\beta,\gamma)}{L_2(\beta,\beta)} \pm \frac{\sqrt{L_2(\alpha,\beta)^2 - L_2(\alpha,\alpha)L_2(\beta,\beta)}\sqrt{L_2(\beta,\gamma)^2 - L_2(\beta,\beta)L_2(\gamma,\gamma)}}{L_2(\beta,\beta)}, \tag{3.10}$$

where in the second step of (3.9) we integrate over the $L_3$ variable. The term $L_3(\alpha,\beta,\gamma)^2$ appearing in the integrand is of course meant to be expressed in terms of the basic $L$-variables according to (2.4). The integration limits in (3.9) reflect exactly the inequalities (I), (II) and (III), since the values for the upper and lower limit of the $L_2(\alpha,\gamma)$-integration correspond to $L_3(\alpha,\beta,\gamma) = 0$ (note that we have $L_3(\alpha,\beta,\gamma)^2 = L_2(\beta,\beta)(L_+ - L_2(\alpha,\gamma))(L_2(\alpha,\gamma) - L_-)$).

We can now transform more of the link variables into loop variables, using formula (3.3). This is best done by enlarging the set of links of fig. 2 step by step to a neighbourhood (see ref. [14] for a definition). However, the gauge-fixing of the links of $C$ used to obtain the measure (3.9) cannot be employed for neighbouring unit cells without creating closed loops of gauge-fixed links (which is not allowed). If this were possible, the full loop variable measure would just be a product of the six-dimensional measure for $C$ over all unit cells of the lattice. Clearly this can only be true for the limit of the theory in which individual unit cells completely decouple. To determine the exact form of this coupling between unit cells in $d$ dimensions is an important next step in our approach, which is currently under investigation.

## 4. The 1×1-lattice

The simplest, but nevertheless instructive example is that of the periodic $1 \times 1$-lattice, consisting of just two links, with associated holonomy matrices $U_\alpha$ and $U_\beta$. We have three independent degrees of freedom and therefore three

$L$-variables [14], which in terms of $U_\alpha$ and $U_\beta$ read

$$
\begin{aligned}
L_1 &= \tfrac{1}{2}\,\mathrm{Tr}\, U_\alpha U_\beta U_\alpha^{-1} U_\beta^{-1}\,,\\
L_2 &= \tfrac{1}{4}\,(\mathrm{Tr}\, U_\beta U_\alpha U_\beta^{-1} U_\alpha U_\beta U_\alpha^{-1} U_\beta^{-1} U_\alpha^{-1} - \mathrm{Tr}\, U_\beta U_\alpha U_\alpha U_\beta^{-1} U_\alpha^{-1} U_\alpha^{-1})\,,\\
L_2' &= \tfrac{1}{4}\,(\mathrm{Tr}\, U_\alpha U_\beta U_\alpha^{-1} U_\beta U_\alpha U_\beta^{-1} U_\alpha^{-1} U_\beta^{-1} - \mathrm{Tr}\, U_\alpha U_\beta U_\beta U_\alpha^{-1} U_\beta^{-1} U_\beta^{-1})\,,
\end{aligned}
\tag{4.1}
$$

where again we choose to gauge-fix the holonomy matrices according to (3.1) and (3.2). The transformation from the canonical group measure to the measure in terms of independent $L$-variables can easily be performed by hand, yielding

$$
\int \mathrm{d}M = 4\pi^2 \int_{-1}^{1} \mathrm{d}L_1 \int_{-(1-L_1^2)}^{1-L_1^2} \mathrm{d}L_2 \int_{-(1-L_1^2)}^{-(1-L_1^2)-\frac{4(L_1-1)(L_1^2+L_2-1)}{4L_1+L_2-L_1^2-3}} \mathrm{d}L_2' \frac{1}{\sqrt{X}}\,, \tag{4.2}
$$

where $X$ is polynomial in $L$,

$$
\begin{aligned}
X = {}& (L_1^2 - L_2 - 1)(L_1^2 - L_2' - 1)\\
&\times \Big(8(L_1-1)^3 + [4(L_1-1) - (L_1^2 - L_2 - 1)][4(L_1-1) - (L_1^2 - L_2' - 1)]\Big)\,.
\end{aligned}
\tag{4.3}
$$

Two features of (4.2) are noteworthy: (i) the measure factor $1/\sqrt{X}$ couples all three $L$-variables in a non-trivial fashion; (ii) the range of integration of $L_2$ is as expected from inequality (II), however, the analogous range of integration of $L_2'$ is strictly smaller. This may be due to the small size and the periodicity of the lattice, which leads to multiple occurrences of the link variables in (4.1).

The partition function $P(\beta)$ (with $\beta = g^{-2}$) for the $1\times 1$-lattice is obtained by averaging the exponentiated Wilson action over all possible configurations. Since the action does not depend on the $L_2$ variables, we can integrate them out and are left with the integral

$$
P(\beta) = 4\pi^3 \int_{-1}^{1} \mathrm{d}L_1\, (\tfrac{1}{2}\pi + \arcsin L_1)\, \mathrm{e}^{-\beta(1-L_1)} = \frac{4\pi^4}{\beta}(1 - \mathrm{e}^{-\beta} I_0(\beta)\,)\,. \tag{4.4}
$$

The expectation value of the one-plaquette variable $L_1$ is found to be

$$
\langle L_1 \rangle = -\frac{1}{\beta} + \frac{1 - I_1(\beta)\mathrm{e}^{-\beta}}{1 - I_0(\beta)\mathrm{e}^{-\beta}}\,, \tag{4.5}
$$

where $I_0$ and $I_1$ denote the Bessel functions of imaginary arguments [20].

The example of the $1\times 1$-lattice also illustrates the relation between the independent lattice variables and the Polyakov loops mentioned in sect. 2. In the present case the basic non-contractible loops are simply the links $\alpha$ and $\beta$, with endpoints identified because of the periodic boundary conditions. The corresponding Polyakov loops are the variables $\mathrm{Tr}\, U_\alpha$ and $\mathrm{Tr}\, U_\beta$. It is now easy to show that from the knowledge of the independent loop variables $L_1$, $L_2$ and $L'_2$ we can only obtain the *square* of a Polyakov loop,

$$(\mathrm{Tr}\, U_\alpha)^2 = \frac{L_2 - L_1^2 + 1}{1 - L_1}, \tag{4.6}$$

and similarly for $(\mathrm{Tr}\, U_\beta)^2$. We conclude that the $L$-variables do not see the two discrete $\mathbb{Z}_2$ degrees of freedom associated with the Polyakov loops, i.e. they form a multiple (4-fold) cover of the physical configuration space $\mathrm{Q}_{\mathrm{red}}$. If one starts from link variables in the adjoint rather than the fundamental representation, it is straightforward to see that the map between the $L$-variables and $\mathrm{Q}'_{\mathrm{red}}$ becomes 1-to-1, since the traces in these representations are related by

$$\mathrm{Tr}_{\mathrm{adj}} U = (\mathrm{Tr}_{\mathrm{fund}} U)^2 - 1. \tag{4.7}$$

With link variables in the adjoint representation, the $\mathbb{Z}_2$ centre symmetry becomes a topological symmetry. Since the gauge group effectively is $\mathrm{SU}(2)/\mathbb{Z}_2$, and the lattice (as a torus) contains $\mathrm{S}^1$'s, the fact that the fundamental group $\pi_1(\mathrm{SU}(2)/\mathbb{Z}_2) = \mathbb{Z}_2$ allows one to retrieve this symmetry [15].

## 5. Lattice loop equations

Since the explicit form of the gauge-covariant link variables $U_l$ in terms of the independent loop variables $L$ is (up to now) not known for general hypercubic lattices, there is no straightforward way of rewriting well-known expressions from the standard covariant formulation as functions of the $L$-variables. In any case it is an interesting question whether there is an *independent* way of deriving classical Yang–Mills equations in the loop formulation, without making reference to the usual connection formulation. Neither for gauge nor for gravity theories (in the Ashtekar formulation; see ref. [21] for a review) an action principle for gauge-invariant loop variables has been established, be it in the continuum or some discretized version of the theory.

We will in the following propose a consistent way of deriving equations of motion in terms of independent loop variables on the hypercubic lattice, by finding the stationary points of the Wilson action. The special feature of our formulation in terms of gauge-invariant variables is the fact that the action is linear in the basic variables, and naively writing down Euler–Lagrange equations (corresponding to linear variations in the $L_i$ variables) leads to inconsistencies. We show that in order to arrive at consistent equations of motion and the vacuum solution (which has to be present), one has to take into account that infinitesimal variations are restricted by the inequalities between the configuration space variables. This is a consequence of working in the non-linear space $Q_{red}$ of physical variables. A similar method for obtaining equations of motion will have to be employed in an analogous continuum loop formulation of Yang–Mills theory, if one uses a Wilson-type action. Our derivation is the analogue for gauge-invariant variables of Polyakov's [4] derivation of lattice equations for gauge theory. In the process, we introduce a realization of the $L$-variables in terms of a space of angular variables, which exhibit the non-linear structure of the physical configuration space and the geometrical origin of the inequalities (I)–(III), and are likely to play a role in the hamiltonian loop formulation.

We begin by discussing the case of two (euclidean) space-time dimensions. Starting point is the familiar Wilson action

$$S_W[U_l] = -\frac{1}{2g^2}\sum_{\gamma_{ij}}(2 - \operatorname{Tr} U_{\gamma_{ij}}), \tag{5.1}$$

where the sum replaces the integration over space-time and is taken over all plaquettes $\gamma$ consisting of four contiguous oriented links $l_i$, $\gamma = l_1 \cdot l_2 \cdot l_3 \cdot l_4$. Recall in two dimensions there are $3N^2$ independent loop variables, three associated with each unit cell in position $(ij)$ [14]: $L_1(\gamma_{ij})$, $L_2(\gamma_{ij}, \gamma_{i,j+1})$ and $L_2(\gamma_{ij}, \gamma_{i+1,j})$ (with $\gamma_{ij}$ denoting the corresponding plaquette loop). Note that the action is already a function of the subset of independent loop variables $\{L_1(\gamma_{ij}) = \frac{1}{2}\operatorname{Tr} U_{\gamma_{ij}}\}$, whence we may rewrite it in the form

$$S'_W[L] = -\frac{1}{g^2}\sum_{\gamma_{ij}}(1 - L_1(\gamma_{ij})). \tag{5.2}$$

Since the action (5.2) is just a linear function of the basic $L_1$ variables, it is a priori not clear how non-trivial equations of motion can be obtained from it. The resolution of this apparent contradiction is the observation that the $L$-variables are constrained by a set of inequalities (only type (I) and (II) are relevant in two dimensions), and that infinitesimal variations of (5.2) must be chosen in a way which respects these inequalities. We have to find $3N^2$ independent variations on the physical configuration space $Q_{red}$ parametrized by the loop variables. Obviously an additive variation like $L_1(\gamma) \to L_1(\gamma) + \epsilon$ violates (I) and hence is not allowed. It turns out that typical admissible variations are non-linear in the basic variables.

The following is based on the assumption that the inequalities (I) and (II) do indeed exhaust the available range of the $L$-variables. This is certainly true for sufficiently small neighbourhoods of plaquettes on sufficiently big ($N > 2$) lattices. However, there may be further restrictions on the $L$-variables, which are not immediately obvious from the derivation in sect. 2, for example, as a result of the periodic boundary conditions. This occurred in the case of the $1 \times 1$-lattice described in the previous section. What happens in the general case we will only be able to answer after calculations for bigger lattices have been performed, but this will presumably not invalidate the general method of obtaining the equations of motion.

Let us establish some notation first. We will say that a point (given by the values of $3N^2$ $L$-variables) lies "on the boundary" of configuration space if some or all of the inequalitites (II) are strict equalities; otherwise it will be called an interior point. Note that points where some of the expressions (I) become equalities are automatically boundary points.

There is a realization of the $L$-variables in terms of a set of angular variables such that the inequalities (I) and (II) are satisfied automatically. Set

$$\begin{aligned} L_1(\gamma_{ij}) &= \cos\Theta_{ij}, \\ L_2(\gamma_{ij},\gamma_{i,j+1}) &= \sqrt{\sin^2\Theta_{ij}\sin^2\Theta_{i,j+1}}\cos\phi_{ij;i,j+1}, \\ L_2(\gamma_{ij},\gamma_{i+1,j}) &= \sqrt{\sin^2\Theta_{ij}\sin^2\Theta_{i+1,j}}\cos\phi_{ij;i+1,j}, \end{aligned} \tag{5.3}$$

for $N^2$ angles $\Theta_{ij} \in [0,2\pi]$ and $2N^2$ angles $\phi_{ij;kl} \in [0,2\pi]$. Infinitesimal rotations in terms of these angular variables induce infinitesimal transformations on the $L$-variables which by construction respect the inequalities (I) and (II). The first kind of infinitesimal transformation (coming from a variation $\Theta_{ij} \to \Theta_{ij} + \epsilon$) acts non-trivially on an $L_1$ variable $L_1(\gamma_{ij})$ and the four $L_2$ variables which have the plaquette loop $\gamma_{ij}$ as one of their arguments. For $\epsilon \ll 1$, define

$$\begin{aligned} L_1(\gamma_{ij}) &\to L_1(\gamma_{ij}) - \epsilon\sqrt{1 - L_1(\gamma_{ij})^2}, \\ L_2(\gamma_{ij},\gamma') &\to L_2(\gamma_{ij},\gamma') + \epsilon L_2(\gamma_{ij},\gamma')\frac{L_1(\gamma_{ij})}{\sqrt{1 - L_1(\gamma_{ij})^2}}, \end{aligned} \tag{5.4}$$

where $\gamma'$ denotes any of the four plaquettes adjacent to $\gamma_{ij}$. There are $N^2$ transformations of this type. The second kind of transformation (coming from a variation $\phi \to \phi + \epsilon$) acts non-trivially on a single $L_2$ variable $L_2(\gamma,\gamma')$ in such a way that $||L_2(\gamma,\gamma')||$ always remains smaller or equal to $\sqrt{L_2(\gamma,\gamma)L_2(\gamma',\gamma')}$. For $\epsilon \ll 1$, we have

$$L_2(\gamma,\gamma') \to L_2(\gamma,\gamma') - \epsilon\sqrt{L_2(\gamma,\gamma)L_2(\gamma',\gamma') - L_2(\gamma,\gamma')^2}, \tag{5.5}$$

with all other $L$-variables remaining unchanged. Obviously there are $2N^2$ such infinitesimal transformations, one for each $L_2$ variable. We observe that in two dimensions the $L_2$ variables completely decouple from the classical equations of motion: the Euler–Lagrange equations corresponding to (5.4) vanish identically and we are left with $N^2$ trivial equations coming from (5.2), (5.3). They simply are given by

$$\sqrt{1 - L_1(\gamma_{ij})^2} = 0, \tag{5.6}$$

which has as solutions

$$L_1(\gamma_{ij}) = \pm 1, \tag{5.7}$$

with an overall minus sign corresponding to a maximum and an overall plus sign corresponding to a minimum of the action $S'_W$, in fact, the vacuum solution (given by $U_l = \mathbb{1}$ for all link holonomies).

The $3N^2$ infinitesimal variations are independent at each interior point of $Q_{red}$ and have maximal span everywhere on the boundary (which has dimension lower than $3N^2$, depending on how many of the expressions (I) and (II) are strict equalities). Note furthermore that any solution to the equations of motion has to lie on the boundary, because the action (5.2) is linear in the basic $L$-variables.

The situation in three dimensions is somewhat more interesting, due to the appearance of the inequality (III). Let $L_1(\alpha)$, $L_1(\beta)$, $L_1(\gamma)$, $L_2(\alpha,\beta)$, $L_2(\alpha,\gamma)$ and $L_2(\beta,\gamma)$ be the $L$-variables associated with a particular unit cell on the three-dimensional lattice (cf. (3.6)). Again we can employ the realization (5.3) with a set of angles $\Theta_\alpha$, $\Theta_\beta$, $\Theta_\gamma$, $\phi_{\alpha\beta}$, $\phi_{\alpha\gamma}$ and $\phi_{\beta\gamma}$, but now there is an additional constraint on these data, coming from inequality (III), namely

$$\sin^2\Theta_\alpha \sin^2\Theta_\beta \sin^2\Theta_\gamma (1 - \cos^2\phi_{\alpha\beta} - \cos^2\phi_{\alpha\gamma} - \cos^2\phi_{\beta\gamma} + 2\cos\phi_{\alpha\beta}\cos\phi_{\alpha\gamma}\cos\phi_{\beta\gamma}) \geqslant 0. \tag{5.8}$$

This is always fulfilled if we choose the three $\phi$-angles in such a way that the inequality

$$\cos(\phi_{\alpha\beta} - \phi_{\alpha\gamma}) \leqslant \cos\phi_{\beta\gamma} \leqslant \cos(\phi_{\alpha\beta} + \phi_{\alpha\gamma},, \tag{5.9}$$

or a cyclic permutation of (5.9) with respect to $(\alpha,\beta,\gamma)$ holds. The analogue of transformation (5.4),

$$\begin{aligned} L_1(\alpha) &\to L_1(\alpha) - \epsilon\sqrt{1 - L_1(\alpha)^2}, \\ L_2(\alpha,\beta) &\to L_2(\alpha,\beta) + \epsilon L_2(\alpha,\beta)\frac{L_1(\alpha)}{\sqrt{1 - L_1(\alpha)^2}}, \\ L_2(\alpha,\gamma) &\to L_2(\alpha,\gamma) + \epsilon L_2(\alpha,\beta)\frac{L_1(\alpha)}{\sqrt{1 - L_1(\alpha)^2}}, \end{aligned} \tag{5.10}$$

is again compatible with all the inequalities, and there are similar transformations associated with the other plaquette loops, $\alpha$ and $\beta$. In total, there are $3N^3$ such infinitesimal variations on the lattice.

In three dimensions, we define the boundary of the physical configuration space to consist of all points where (III) is a strict equality for some or all of the $L$-variables. One way of defining the remaining infinitesimal variations is as

$$L_2(\alpha,\beta) \to L_2(\alpha,\beta) - \epsilon\sqrt{L_3(\alpha,\beta,\gamma)^2}, \tag{5.11}$$

(where again $L_3$ is short-hand for the left-hand side of (2.4),) and leaving all other $L$-variables unchanged. There are two other infinitesimal variations, with $L_2(\alpha,\beta)$ substituted by $L_2(\alpha,\gamma)$ and $L_2(\beta,\gamma)$ respectively, i.e. a total of $3N^3$ such transformations for the whole lattice.

It is straightforward to check that the infinitesimal variations (5.10) and (5.11) (six variations per unit cell) are independent everywhere in the interior of $Q_{red}$. Since the Wilson action in three dimensions is linear in $L_1$, this again implies all the solutions to the equations of motion have to lie on the boundary of $Q_{red}$. Note that the $\mathbb{Z}_2$-ambiguity related to the sign of $L_3(\alpha,\beta,\gamma)$ (in the description of configuration space in terms of the independent $L_1$ and $L_2$ variables only) does not play a role here. The two sectors associated with $L_3(\alpha,\beta,\gamma) > 0$ and $L_3(\alpha,\beta,\gamma) < 0$ meet along the boundary defined by $L_3(\alpha,\beta,\gamma) = 0$.

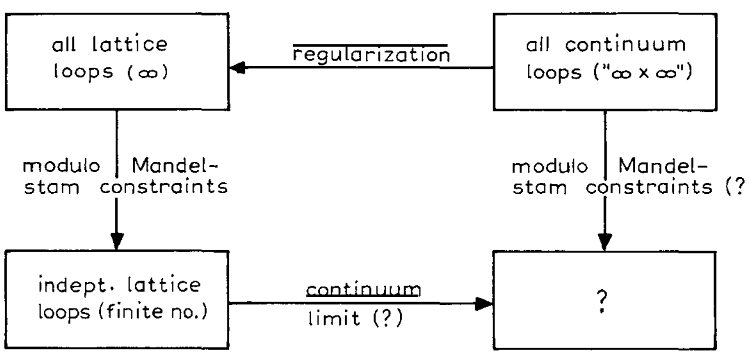
Unlike in the two-dimensional case, the vector fields associated with (5.10) and (5.11) do not span the tangent space of $Q_{red}$ everywhere on the boundary. However, even if we modified (5.11) appropriately to achieve maximal span on the boundary, the equations of motion coming from (5.10) will again be given by (5.6), leading to the solutions (5.7), which imply $L_2 \equiv 0$ for all of the $L_2$ variables.

Several comments on these results are in order: (i) The Wilson action (5.1) is only the simplest possible choice, and there are many other actions coinciding with the usual Yang–Mills action in the continuum limit. Several authors have suggested modified forms of the lattice action, to improve the lattice cut-off dependence of certain physical quantities (see for example refs. [22,23]). In our language this would amount to substituting the Wilson lagrangian by some higher-order polynomial of the $L_1$ and $L_2$ variables. For these cases the procedure outlined above may lead to non-trivial classical solutions lying in the interior of $Q_{red}$. With the help of our explicit set of independent traced holonomies, it may also be possible to systematize the search for such an improved SU(2) lattice action. (ii) The choice of independent variations on configuration space is closely linked to the issue of finding appropriate canonically conjugate momenta when going to the hamiltonian version of the theory. In this context the realization of $Q_{red}$ in terms of the angular variables $(\Theta,\phi)$ introduced above may play an important role. This could also lead to new insights into the structure of natural differential operators on the loop space in the continuum theory. (iii) Although the $L_2$ variables "decouple" from the classical lattice theory, they will of course contribute non-trivially in the quantum theory, for example, by coupling neighbouring unit cells via a non-trivial measure in terms of $L$-variables, as mentioned in sect. 3 above. The case of four dimensions does not contain any qualitatively new results and will not be described here.

## 6. Conclusions

Let us summarize the results obtained so far in the lattice theory: starting from an *infinite* set of lattice loops (and their associated traced holonomies) and an *infinite* number of Mandelstam constraints, we could eliminate all but a finite set of loop variables, describing the physical configuration space of the lattice theory. For a corresponding continuum loop formulation the result seems to imply that in order to extract relevant physical information, it suffices to look at the sector of *small* loops and their associated traced holonomies.

However, it is not even in principle clear how, independently of the lattice formulation, such a statement could be arrived at in the continuum theory: the solution of the Mandelstam constraints on the hypercubic lattice makes crucial use of the existence of a smallest loop size (of loops going around just one plaquette), a concept that a priori does not exist in the continuum. It seems that in the continuum loop approach to Yang–Mills theory one has to deal with an additional infinity coming from the Mandelstam constraints, on top of the usual field theoretic infinities. Rephrasing the problem, it is not at all obvious how the following diagram, describing the classical loop kinematics, could be made commutative.




One way of getting a step closer toward commutativity would be to *postulate* the existence of a smallest loop size or length scale also in the continuum. In fact, there is a generic problem in loop approaches in that operators or other structures, which are well-defined as long as the loops have finite size, become singular or collapse when their loop arguments are shrunk to points. Such problems are of course aggrevated in the corresponding quantum theories. This suggests that in a pure loop approach one may have to reconsider the structure of the continuum limit, with some basic non-locality remaining. This way some of the lattice results, like the ones described above, would have a much more direct translation into the continuum theory. For example, one may consider a hybrid lattice-continuum description, where one keeps space-time continuous, but selects a discrete set of, say, rectangular loops per space-time point (three in two dimensions, six in three dimensions etc.), and tries to rewrite Yang–Mills theory in terms of their associated traced holonomies.

A still outstanding problem in path-dependent approaches is the need for an intrinsically path-dependent perturbation theory. If there is a fundamental *non*-equivalence between the path-dependent and the usual local formulation of Yang–Mills theory at the quantum level, (a hope that is behind most of the current non-local approaches,) resorting to a local perturbation theory in terms of the gauge potentials $A_\mu$ clearly defeats this purpose. Another known alternative in SU$(N)$ gauge theory, the $1/N$-expansion, cannot be applied straightforwardly in our approach, which depends strongly on the dimension $N$ of the group matrices. We hope that the results derived in this paper will help in developing a more rigorous formulation of Yang–Mills theory in terms of loop variables.

I am grateful to P. van Baal for his detailed comments on the topology of lattice gauge fields.

## References

[1] I. Bialynicki-Birula, Bull. Acad. Polon. Sci. 11 (1963) 135
[2] S. Mandelstam, Phys. Rev. 175 (1968) 1580
[3] A.M. Polyakov, Phys. Lett. B82 (1979) 247
[4] A.M. Polyakov, Gauge fields and strings (Harwood, New York, 1987)
[5] Yu.M. Makeenko and A.A. Migdal, Nucl. Phys. B188 (1981) 269
[6] A.A. Migdal, Phys. Rep. 102 (1983) 199
[7] R. Giles, Phys. Rev. D24 (1981) 2160
[8] R. Loll, Loop approaches to gauge field theory, to appear in Teor. Mat. Fiz.
[9] H. Fort and R. Gambini, Phys. Rev. D44 (1991) 1257
[10] W. Furmanski and A. Kolawa, Nucl. Phys. B291 (1987) 594
[11] R. Gambini, L. Leal and A. Trias, Phys. Rev. D39 (1989) 3127
[12] C. Rovelli and L. Smolin, Loop representation for lattice gauge theory, preprint Pittsburgh and Syracuse (1990)
[13] B. Brügmann, Phys. Rev. D43 (1991) 566
[14] R. Loll, Nucl. Phys. B368 (1992) 121
[15] P. van Baal, Commun. Math. Phys. 85 (1982) 529 and private communication
[16] A. Das, M. Kaku and P. Townsend, Nucl. Phys. B149 (1979) 109
[17] J. Goldstone and R. Jackiw, Phys. Lett. B74 (1978) 81
[18] J. Kogut, Rev. Mod. Phys. 55 (1983) 775
[19] M. Lüscher, Commun. Math. Phys. 85 (1982) 39
[20] I.S. Gradsteyn and I.M. Ryzhik, Table of integrals, series, and products, corrected and enlarged edition (Academic Press, San Diego, 1980)
[21] C. Rovelli, Class. Quantum Grav. 8 (1991) 1613
[22] P. Weisz, Nucl. Phys. B212 (1983) 1
[23] B. Berg, A. Billoire, S. Meyer and C. Panagiotakopoulos, Commun. Math. Phys. 97 (1985) 31