Yang–Mills theory in terms of gauge invariant dual variables

Dmitri Diakonov$^{\diamond *}$

$^{\diamond}$ *NORDITA, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark*

* *St. Petersburg Nuclear Physics Institute, Gatchina 188 300, Russia*

Abstract

Quantum Yang–Mills theory and the Wilson loop can be rewritten identically in terms of local gauge-invariant variables being directly related to the metric of the dual space. In this formulation, one reveals a hidden high local symmetry of the Yang–Mills theory, which mixes up fields with spins up to $J = N$ for the $SU(N)$ gauge group. In the simplest case of the $SU(2)$ group the dual space seems to tend to the de Sitter space in the infrared region. This observation suggests a new mechanism of gauge-invariant mass generation in the Yang–Mills theory [1].

1   Introduction

Ever since the formulation of Quantum Chromodynamics and the realization that only gauge-invariant operators are the observables, there have been attempts to reformulate the theory itself not in terms of the Yang–Mills potentials but rather in terms of some gauge-invariant variables. However, usually it is difficult to incorporate matter in such approaches as it couples not to gauge-invariant variables but to the gauge potentials A_μ.

The simplest but typical matter is the probe source, or the Wilson loop. It is given by a path-ordered exponent of A_μ, and it is difficult to present it in terms of any gauge-invariant variables. At some point it has been suggested that Wilson loops themselves are the only

---

[1]Invited talk at *Confinement-5*, Lago di Garda, 10-14 Sep 2002.

reasonable gauge-invariant variables, and the theory should be reformulated in terms of loop dynamics [1, 2]. Because Wilson loops are non-local objects there have been no decisive success on this way, despite enormous efforts taken in twenty years.

Recently, it became clear that it is not only possible to reformulate exactly the YM theory (together with the Wilson loop) in terms of *local* gauge-invariant variables, but that the resulting theory is beautiful and suggestive. It reminds quantum gravity theory, where the metric and curvature refer to the *dual* YM space (below I specify what precisely does it mean). It becomes possible to discuss the long-standing but still unresolved questions like the mass generation and the area behaviour of the Wilson loop directly in a gauge-invariant, *i.e.* objective way. Also, the common belief is that, at least at large number of colours N, the YM theory is equivalent to some version of string theory. String theory is formulated in terms of gauge-invariant although non-local variables. Recent attempts to justify this equivalence from the string side have been described by A. Polyakov [3]. If one wishes to derive string theory directly from the YM Lagrangian, the first step is to rewrite the theory in terms of gauge-invariant variables. I shall show below that the YM theory possesses an exciting new symmetry mixing states with different spins, which makes the equivalence with string theory rather likely.

2   First order formalism

2.1   Four dimensions

An economic way to introduce gauge-invariant variables is via the so-called first order formalism. The $4d$ YM partition function can be identically rewritten with the help of an additional Gaussian integration over dual field strength variables [4, 5]:

$$
\begin{aligned}
Z_{4\mathrm{d}} &= \int DA_\mu \, \exp \int d^4x \, \left(-\frac{1}{2g_4^2} \mathrm{Tr}\, F_{\mu\nu} F_{\mu\nu} \right) \\
&= \int DA_\mu \, DG_{\mu\nu} \, \exp \int d^4x \left(-\frac{g_4^2}{2} \mathrm{Tr}\, G_{\mu\nu} G_{\mu\nu} + \frac{i}{2}\, \epsilon^{\alpha\beta\mu\nu} \, \mathrm{Tr}\, G_{\alpha\beta} F_{\mu\nu} \right)
\end{aligned} \tag{1}
$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu A_\nu]$ is the standard Yang–Mills field strength and $\epsilon^{\alpha\beta\mu\nu}$ is the antisymmetric tensor. To be specific, the gauge group is $SU(N)$ with N^2-1 generators t^a, $\mathrm{Tr}\, t^a t^b = \delta^{ab}/2$. Eq. (1) is called the 1st order formalism.

Both terms in eq. (1) are invariant under (N^2-1)-function gauge transformation,

$$
\begin{cases}
\delta A_\mu &= [D_\mu \, \alpha], \\
\delta G_{\mu\nu} &= [G_{\mu\nu} \, \alpha]
\end{cases} \tag{2}
$$

where $D_\mu = \partial_\mu - i A_\mu^a t^a$ is the Yang–Mills covariant derivative, $[D_\mu D_\nu] = -iF_{\mu\nu}$.

Owing to the Bianchi identity, $\epsilon^{\mu\nu\rho\sigma}\,[D_\nu F_{\rho\sigma}] = 0$, the second (mixed) term in eq. (1) is, in addition, invariant under the $4(N^2-1)$-function 'dual' gauge transformation,

$$
\begin{cases}
\delta A_\mu &= 0, \\
\delta G_{\mu\nu} &= [D_\mu \beta_\nu] - [D_\nu \beta_\mu].
\end{cases} \tag{3}
$$

Taking a particular combination of the functions in eqs. (2,3),

$$\alpha = v^\mu A_\mu, \qquad \beta_\mu = v^\lambda G_{\lambda\mu}, \tag{4}$$

leads to the 4-function transformation

$$\delta G_{\mu\nu} = -G_{\lambda\nu}\, \partial_\mu v^\lambda - G_{\mu\lambda}\, \partial_\nu v^\lambda - \partial_\lambda G_{\mu\nu}\, v^\lambda, \tag{5}$$

being the known variation of a (covariant) tensor under infinitesimal general coordinate transformation, $x^\mu \to x^\mu + v^\mu(x)$, also called the diffeomorphism. Therefore, the 'mixed' term is diffeomorphism-invariant, and is known as BF gravity [2]. It defines a topological field theory of the Schwarz type [6]. Moreover, it is invariant not under four but as much as $4(N^2-1)$ local transformations; four diffeomorphisms are but their small subset. We shall see later on that the additional local transformations mix up fields with different spins.

In the first order formalism (1), the YM potential A_μ enters linearly and quadratically. Therefore, on can now perform the Gaussian integration over A_μ: the quadratic form for A_μ is, generally, non-degenerate and contains no derivatives. Thus, one has just a product of Gaussian integrals at each point, and it is not too difficult to perform the A_μ integration explicitly [4, 5]. The point is to write the result in a nice way such that, for example, the diffeo-invariance is explicit. This has been done in Ref. [7], however the result there is more than one page long, and the invariance under $4(N^2-1)$-function transformation (3) has been neither revealed nor discussed. It has been achieved in a compact way in Ref. [8] which we review in the next section.

2.2 Three dimensions

We shall be also interested in the $3d$ YM theory: it is also believed to possess the area law for large Wilson loops and an exponential decrease of correlators of small Wilson loops at large separations. The theory is super-renormalizable; the mass gaps and string tension are proportional to the gauge coupling g_3^2 (having the dimension of mass) in the appropriate power.

In the $3d$ case one can also use the Gaussian trick to rewrite the partition function in the first order formalism. In this case the dual field strength is a 3-vector, call it G_i^a (the Latin indices go from 1 to 3).

$$
\begin{aligned}
Z_{3\mathrm{d}} &= \int DA_i^a \, \exp\left(-\frac{1}{4g_3^2}\int d^3x\, F_{ij}^a(A)F_{ij}^a(A)\right) \\
&= \int DG_i^a\, DA_i^a\, \exp\int d^3x \left[-\frac{g_3^2}{2}G_i^a G_i^a + \frac{i}{2}\epsilon^{ijk}F_{ij}^a(A)G_k^a\right].
\end{aligned}
\tag{6}
$$

As in the $4d$ case, the Bianchi identity, $\epsilon^{ijk}D_i^{ab}F_{jk}^b = 0$, ensures that the second (mixed) term in eq. (6) is invariant under local (N^2-1)-function dual gauge transformation

$$\begin{cases} \delta A_i^a &= \quad 0, \\ \delta G_i^a &= \quad D_i^{ab}(A)\,\beta^b \end{cases} \tag{7}$$

---

[2]With our notations it would be more appropriate to call it 'GF gravity' but we follow the tradition.

Again, the invariance under ordinary gauge transformation together with the dual one guarantees that, if one integrates out the YM potentials A_i, the resulting action will be invariant under local $(N^2 - 1)$-function transformations, three of which are the diffeomorphisms. For $SU(2)$ that is the only invariance but at $N > 2$ the invariance is much wider than that of the general relativity.

3 Yang–Mills theory as quantum gravity with 'æther'

In this section we give the results of integrating out the YM potential A_μ from the first-order formalism's eqs. (1,6). The goal is to obtain the result in terms of gauge-invariant combinations of the dual field strength $G_{\mu\nu}$, G_i and to see that the second (mixed) term has the claimed large local invariance.

3.1 Three dimensions, $SU(2)$

We start with the simplest case of the $SU(2)$ gauge group in $3d$. In this case one can identify the dual field strength $G_i^a = e_i^a$ with a covariant *dreibein* [9]. Integrating out A_i^a in eq. (6) one obtains, identically [7, 10, 11, 12]

$$Z_{3\mathrm{d}} = \int Dg_{ij}\, g^{-\frac{5}{4}} \exp \int d^3x \left[-\frac{g_3^2}{2}\, g_{ii} + \frac{i}{2} \sqrt{g} R(g) \right].$$ (8)

where $g_{ij} = e_i^a e_j^a$ is the covariant metric tensor of the dual space constructed, as we see, from the dual field strength. Knowing the metric g_{ij} one can build the Christoffel symbol Γ^i_{jk}, the Riemann tensor R^i_{jkl}, the Ricci tensor R_{ij} and finally the scalar curvature R according to the general formulae of differential geometry, see *e.g.* [14]; g is the determinant of g_{ij}.

The second term in the action (8) is the familiar Einstein–Hilbert action in $3d$; notice that it comes with a purely imaginary 'Newton constant'. The fact that the second term is invariant under the diffeomorphisms of the dual space is the consequence of the invariance of the mixed term in the first order formalism under dual gauge transformations. The first term in the action, g_{ii}, is just a rewriting of the first term from eq. (6); it is not invariant under dual gauge transformations and hence it is not invariant under diffeomorphisms. We call it the 'æther term' as it reminds the coupling of gravity to matter with the stress-energy tensor $T^{ij} = \delta^{ij}$ which is isotropic and homogeneous but makes one system of coordinates preferable. It is this æther term which distinguishes the YM theory from pure Einstein gravity. In $3d$ Einstein's gravity is a non-propagating topological field theory [13].

There are many alternative ways how to parametrize a curved space; defining the metric g_{ij} is only one of them, and not the best in this case. One can define a curved d-dimensional space by (locally) embedding it into a flat $D = d(d+1)/2$-dimensional space. In this case, the flat space is $6d$. 6 functions $w^\alpha(x^1, x^2, x^3)$, $\alpha = 1...6$, called the *external coordinates* define the embedding of a generic $3d$ curved manifold into a flat $6d$ space. The metric is induced by this embedding:

$$g_{ij} = \partial_i w^\alpha \partial_j w^\alpha. \tag{9}$$

All differential geometry's quantities including the scalar curvature R can be expressed through the six functions $w^\alpha(x)$ [11]. Invariance under diffeomorphisms is invariance under the re-parametrization: $w^\alpha(x) \to w'^\alpha(x)$. In terms of the external coordinates the YM partition function takes the form [11, 12]

$$
\begin{aligned}
Z_{3\mathrm{d}} &= \int Dw^\alpha(x) \det(\partial_i \partial_j w^\alpha) \, g(w)^{-\frac{5}{4}} \\
&\quad \cdot \exp \int d^3 x \left[-\frac{g_3^2}{2} \partial_i w^\alpha \partial_i w^\alpha + \frac{i}{2} \sqrt{g(w)} \, R(w) \right],
\end{aligned} \tag{10}
$$

so that the æther term becomes just the kinetic energy of 6 massless and gauge-invariant fields; however the curvature term brings in the interaction.

In the original formulation of the $SU(2)$ YM theory in $3d$ there were 9 degrees of freedom (dof's) in the gauge potentials A_i^a out of which 3 were gauge transformations. The remarkable achievement of eqs. (8,10) is that the partition function is presented in terms of exactly 6 gauge-invariant variables: g_{ij} or w^α. Averages of Wilson loops, since they are gauge-invariant, can be also presented in terms of the gauge-invariant parametrization of the dual space. It turns out [12, 14] that Wilson loops become, in the 'gravity' formulation, *parallel transporters* in the curved dual space!

3.2 Three dimensions, general $SU(N)$

In the $3d$ $SU(N)$ YM theory there are $3(N^2 - 1)$ dof's in the original formulation via the gauge potentials A_i^a, out of which $N^2 - 1$ are gauge transformations, therefore, the number of gauge-invariant dof's is $2(N^2 - 1)$.

As in the $SU(2)$ case, we wish to integrate out the YM potentials A_i^a and express the result in terms of gauge-invariant combinations of the dual field strength G_i^a. In the $SU(2)$ case the $2 \cdot 3 = 6$ gauge-invariant variables are quadratic combinations, $g_{ij} = G_i^a G_j^a$; the metric tensor can be decomposed into spin 0 (1 dof) and spin 2 (5 dof's) fields, $1 + 5 = 6$.

For $SU(3)$ one needs $2 \cdot 8 = 16$ gauge-invariant dof's; these are the 6 components of the metric tensor g_{ij}, plus 10 components of spin 1 (3 dof's) and spin 3 (7 dof's) fields, put together into a symmetric rank-3 tensor $h_{ijk} = \frac{1}{3}\mathrm{Tr}\{G_i G_j G_k\}$ where $G_i = G_i^a t^a$ and $\{\}$ means full symmetrization. In $SU(2)$ h_{ijk} is automatically zero. In $SU(4)$ one would need a rank-4 symmetric combination of G_i's which is automatically zero in $SU(3)$, and so on [8].

The general pattern is that for the $SU(N)$ gauge group one adds new spin N and spin $N-2$ fields to the 'previous' fields of the $SU(N-1)$ group. All in all, one has for the $SU(N)$ two copies of spin $2, 3 \ldots N-2$ and one copy of the 'edge' spins 0, 1, $N-1$ and N; they sum up into the needed $2(N^2 - 1)$ dof's [8]. In the usual formulation, one gets $2(N^2 - 1)$ dof's just because there is a colour index which runs up to $N^2 - 1$, however the presence of a colour index implies the quantity is not gauge-invariant. If one wishes to pass to gauge-invariant variables, it is inevitable that the 'coloured' dof's are 'squeezed' into higher and higher spin fields.

The invariance of the second term in the first-order-formalism action (6) under (N^2-1)-function dual gauge transformation (7) translates into (N^2-1)-function local symmetry which mixes fields with different spins [8]. This exciting new symmetry can be paralleled to that of supergravity where one can locally mix up spin 2 and spin 3/2. However, in our case only boson fields are mixed by symmetry and one can mix up the whole tower of spins up to $J = N$. The transformation is non-linear, though, so there is no contradiction with the Coleman–Mandula theorem.

After integrating out A_i^a, the second term reminds the Einstein–Hilbert action but contains not only the metric tensor g_{ij} but also higher spin fields $h_{ijk} \ldots$. Since $SU(2)$ is a subgroup of any Lie group, this part of the action is diffeomorphism-invariant but for $N > 2$ it is in addition invariant under other local transformations mixing fields with different spins [8]. At $N \to \infty$ an infinite tower of spins are related by symmetry transformation. Since it is the kind of symmetry known in string theory, and some kind of string is expected to be equivalent to the Yang–Mills theory in the large N limit, it is tempting to use this formalism as a starting point for deriving a string from the local field YM theory. The æther term $g_3^2 N\, g_{ii}$ breaks this symmetry. Therefore, one expects that the role of this term is to lift the degeneracy of the otherwise massless fields and to provide the string with a finite slope $\alpha' = (g_3^2 N)^{-2}$.

3.3 Four dimensions, $SU(2)$

In $4d$ the dual field strength is an antisymmetric tensor, not a vector, so to construct the metric tensor one has to be more industrious. The needed metric tensor can be constructed as [7]

$$g_{\mu\nu} = \frac{1}{6} \epsilon^{abc} \frac{\epsilon^{\alpha\beta\rho\sigma}}{2\sqrt{g}}\, G_{\mu\alpha}^a G_{\rho\sigma}^b G_{\beta\nu}^c. \tag{11}$$

However, it is not a convenient variable. A more suitable choice is to present the dual field strength in eq. (1) as [8]

$$G_{\mu\nu}^a = d_i^a\, T_{\mu\nu}^i = d_i^a\, \eta_{AB}^i\, e_\mu^A e_\nu^B, \tag{12}$$

where η_{AB}^i is the 't Hooft symbol projecting the $(1,0) + (0,1)$ representation of the Lorentz $SO(4)$ group into the irreducible $(1,0)$ part. e_μ^A can be called the tetrad, and the metric tensor is, as usually, $g_{\mu\nu} = e_\mu^A e_\nu^A$. There are 16 dof's in the tetrad, however three rotations under one of the $SO(3)$ subgroups of the $SO(4)$ Lorentz group does not enter into the combination (12), therefore, the antisymmetric tensor $T_{\mu\nu}^i$ carries 13 dof's.

The 3×3 tensor d_i^a can be called the triad; it is subject to the normalization constraint $\det d_i^a = 1$ and therefore contains 8 dof's. In fact, the combination (12) is invariant under simultaneous $SO(3)$ rotations of T^i and d_i, therefore the r.h.s. of eq. (12) contains $13+8-3 = 18$ dof's, as does the l.h.s. Thus, eq. (12) is a complete parametrization of $G_{\mu\nu}^a$.

It is now clear how to organize the 15 gauge-invariant variables made of $G_{\mu\nu}^a$. These are the 5 dof's contained in a symmetric 3×3 tensor

$$h_{ij} = d_i^a\, d_j^a, \qquad \det h = 1, \tag{13}$$

6

and 13 dof's of $T^i_{\alpha\beta}$. However, h_{ij} and $T^i_{\alpha\beta}$ will always enter contracted in i,j (as it follows from eq. (12)), so that the dof's associated with the simultaneous $SO(3)$ rotation will drop out. In other words, one can choose h_{ij} to be diagonal and containing only 2 dof's. Together with the 13 dof's of $T^i_{\mu\nu}$ they comprise the needed 15 gauge-invariant degrees of freedom.

After integrating out A_μ from the first-order-formalism partition function (1) and expressing the result through T and h one obtains the $4d$ YM partition function in terms of gauge-invariant variables [8]:

$$Z_{4\mathrm{d}} = \int DhDT\, e^{S_1+S_2}, \tag{14}$$

$$S_1 = -\frac{g_4^2}{4}\int d^4x\, T^i_{\mu\nu}\, h_{ij}\, T^j_{\mu\nu}, \tag{15}$$

$$S_2 = \frac{i}{4}\int d^4x\, \sqrt{g}\, R^j_{\,i\,\mu\nu}\, T^{l\,\mu\nu}\, \epsilon_{jlk}\, h^{ki}, \tag{16}$$

where $R^j_{\,i\,\mu\nu}$ is a 'minor' Riemann tensor, see Ref. [8] for details. We see that eq. (16) is covariant both with respect to Greek and Latin indices: it is the manifestation of the invariance of this part of the action under a 12-function *local* transformation following from the invariance under dual gauge transformations (3). Four of these local transformations are diffeomporphisms; the rest mix up spin 0 and spin 2 fields. It is a 'more general relativity' theory. However, in the particular case when $h_{ij} = \delta_{ij}$ the action S_2 reduces to the usual Einstein–Hilbert action, $\sqrt{g}R$, where R is the standard scalar curvature made of $g_{\mu\nu}$.

I would like to note as an aside that one can think of making quantum gravity renormalizable in the following spirit. Start from the 'more general relativity' theory given by the action S_2. It is not only renormalizable but probably exactly solvable. If, for some reason, the field h_{ij} develops spontaneously a v.e.v. $\sim \delta_{ij}$ one gets, at 'low' energies the Einstein's quantum gravity, but renormalizable in the ultra-violet.

4   Dual transformation on the lattice

The duality transformation to gauge-invariant variables can be done directly on the lattice – see Refs. [16, 11] for $3d$ $SU(2)$, Ref. [17] for $4d$ $SU(2)$ and Ref. [18] for a general Lie group. In Ref. [11] a full circle has been performed: starting from discretizing the $3d$ $SU(2)$ YM theory by the lattice, making the duality transformation, going back to the continuum limit, one explicitly recovers the partition function (10).

5   Perturbation theory

One may wonder how the usual perturbation theory with explicit (colour) gluons at short distances looks like in the gauge-invariant formulation. Perturbation theory means expansion in the coupling $g_{3,4}^2$. The curvature term is $O(1)$ in the coupling, therefore in the leading order it has to vanish! In $3d$ it is just the ordinary scalar curvature which has to vanish; in $4d$ it is a somewhat more complicated action S_2. In $3d$, zero curvature means that the

dual space is flat, which means that it is possible to parametrize it by 3 out of the general 6 external coordinates w^a. [Three functions are just a change of variables, like going from Cartesian to spherical coordinates – it does not make a flat space curvy.]

In the gauge-invariant language, perturbation theory is expansion around the flat dual space.

5.1 Where are gluons?

Let us consider the simplest case of the $3d$ $SU(2)$ theory. If the dual space is flat, the æther term is just the kinetic energy of three massless scalar fields w^a. This is exactly what should be expected: the gauge-invariant content of the pertubation theory is one transversely polarized gluon (times three colours). In $3d$ there is only one transverse polarization which can be described by a scalar field.

In $4d$ case, the gauge-invariant content of the perturbation theory are also transverse gluons, in this case there are two transverse polarizations, times three colours. It can be shown that these degrees indeed show up if one requires the curvature term S_2 to vanish [8].

5.2 Where is the Coulomb force?

To explore the potential energy of two static quarks, one has to study the Wilson loop. The presence of the Wilson loop does not allow the parametrization of the metric tensor g_{ij} by gradients $\partial_i w^\alpha$ as in eq. (9) ; a *curl* must be added:

$$g_{ij} = (\partial_i w^\alpha + B_i)(\partial_j w^\alpha + B_j), \qquad \text{Curl } \mathbf{B} = \mathbf{j}, \tag{17}$$

where $\mathbf{j}$ is the current along the Wilson loop,

$$\mathbf{j} = \int d\tau \, \frac{d\mathbf{x}}{d\tau} \, \delta(\mathbf{x} - \mathbf{x}(\tau)). \tag{18}$$

The leading order of the perturbation theory corresponds to the flat dual space, hence only three w's are nonzero. We thus get from the æther term the Coulomb interaction, plus transverse gluons [15]:

$$
\begin{aligned}
< W > \;\; &= \;\; \int Dw^a \, \exp\left[-\frac{g_3^2}{2} \int d^3x \, (\partial_i w^a + B_i)^2 \right] \\
&= \;\; \det{}^{-\frac{3}{2}}(-\partial^2) \, \exp \frac{g_3^2}{2} \int d^3x \, B_i \, \frac{\delta_{ik}\partial^2 - \partial_i\partial_k}{\partial^2} \, B_k \\
&= \;\; \det{}^{-\frac{3}{2}}(-\partial^2) \, \exp \int d\mathbf{x}d\mathbf{y} \, \mathbf{j}(\mathbf{x}) \, \frac{g_3^2}{8\pi|\mathbf{x}-\mathbf{y}|} \, \mathbf{j}(\mathbf{y}),
\end{aligned} \tag{19}
$$

which is nothing but the Coulomb interaction between segments of the Wilson loop, as it should be in perturbation theory.

8

6   A quick (but wrong) way to get confinement

The aim of rewriting the YM theory in gauge-invariant variables is to get new insights into the nonperturbative regime of the theory. One bonus can be obtained immediately by solving the classical equations. We take the simplest case of the $3d$ $SU(2)$ theory (8) and consider the usual Wilson loop as a source. Varying the action of eq. (8) with respect to the metric we get the classical Einstein's equation modified by the presence of the æther term and the stress-energy tensor corresponding to the Wilson loop:

$$R^{ij} - \frac{1}{2}\,R\,g^{ij} - \frac{ig_3^2}{\sqrt{g}}\,\delta_j^i = iT^{ij}. \tag{20}$$

The æther term breaks the general covariance, of course. This equation can be solved for a large loop lying in a plane, and one gets for the average of the Wilson loop [15]

$$< W > = \exp\left(-\text{Area}\,g_3^2\,\delta(0)\right) \tag{21}$$

i.e. the area law but with an infinitely thin string, because it is a classical calculation. Quantum fluctuations are to smear out the thin string and yield a finite string tension. Nevertheless, it is amusing that in the dual formulation the first thing one gets is the area law!

7   Order-of-magnitude analysis

The reformulation of the YM theory in terms of gauge-invariant dual variables is quite unusual and to get some insight into the ensuing quantum theory, let us first of all analyze the order of magnitude of the fields and of their derivatives. We use the partition function (10). Fluctuations leading to the action much greater than unity are usually cut out.

Let a be some UV cutoff, *e.g.* the lattice spacing. From the first (æther) term we estimate:

$$a^3 g_3^2 (\partial w)^2 \sim 1 \Rightarrow \partial w \sim a^{-1}\,(g_3^2 a)^{-\frac{1}{2}}, \tag{22}$$

hence $\sqrt{g} \sim (\partial w)^3 \sim a^{-3}(g_3^2 a)^{-\frac{3}{2}}$. The second (Einstein–Hilbert) term must be also of the order of unity (or less), otherwise the configuration of the field $w^\alpha(x)$ will be damped by oscillations from nearby fluctuations. We have therefore:

$$a^3 R\sqrt{g} \sim 1 \Rightarrow R \le (g_3^2 a)^{\frac{3}{2}} \to 0 \quad (!) \tag{23}$$

Unlike pure quantum gravity, the æther term requires that quantum fluctuations of the metric should correspond to zero curvature R in the limit of vanishing lattice spacing! [A similar analysis applies to the $4d$ dual theory, with the only difference that the curvature has to vanish not as a power but as one-over-logarithm of the cutoff.] Does it mean that only small fluctuations around flat space are allowed, leading to perturbation theory? Not necessarily. Non-zero values of R are not prohibited provided $R\sqrt{g}$ is a *full derivative.* If so, large values

of $R\sqrt{g}$ cancel between neighbour points, and the oscillating factor, $\exp\left(i\int R\sqrt{g}\right)$, does not damp such metrics.

An example of a non-flat metric with $R\sqrt{g}$ being a full derivative is the (anti) de Sitter S^3 space. It can be parametrized by 4 (out of the general 6) external coordinates w^A lying on a sphere, $\sum_{A=1}^{4} w^A w^A = \frac{6}{R}$. $R\sqrt{g}$ becomes then the winding number of the map $S^3 \mapsto S^3$, which *is* a full derivative [12]. In the next section I show that dual space of constant curvature seems to be, indeed, dynamically preferred.

8 Beyond perturbation theory: de Sitter dual space

Yet another parametrization of the curved dual space is obtained by expressing three external coordinates through the three other and introducing three functions $U^\alpha = U^{1,2,3}(w^1, w^2, w^3)$ such that

$$g_{ij} = \partial_i w^a \partial_j w^b G_{ab}(w), \qquad G_{ab}(w) = \delta_{ab} + \frac{\partial U^\alpha}{\partial w^a}\frac{\partial U^\alpha}{\partial w^b}. \tag{24}$$

[It is similar to describing the $2d$ surface by a function $z(x, y)$]. The æther term becomes now the Lagrangian of a generalized σ model:

$$\mathcal{L} = \partial_i w^a \partial_i w^b G_{ab}(w). \tag{25}$$

In $2 + \epsilon$ dimensions integrating out frequences from μ_1 to μ_2 renormalizes the metric as follows [19]:

$$\frac{dG_{ab}}{d\ln\frac{\mu_2}{\mu_1}} = \epsilon\, G_{ab} - \frac{1}{2\pi}R_{ab} - \frac{1}{8\pi^2}R_{ac}G^{cd}R_{db} - \dots \tag{26}$$

where R_{ab} is the Ricci tensor built from the metric G_{ab} (24). This renormalization-group equation means that the metric gets renormalized as one integrates out high frequences in quantum fluctuations. Putting boldly $\epsilon = 1$ one discovers that the infrared-fixed point (*i.e.* the zero of the r.h.s.) is the de Sitter space with curvature $R = 6\pi$ (to one loop accuracy) or $R = 0.71 \cdot 6\pi$ (in 2 loops), etc. Therefore, at low momenta the theory is described by the $O(n)$ σ-model with $n = 4$, given by the partition function

$$
\begin{aligned}
Z &= \int Dw^A \delta\left(w^2 - \frac{6}{R}\right) \exp\left(-\int d^3x\, \frac{g_3^2}{2}\partial_i w^A \partial_i w^A\right) \\
&= \int D\lambda \int Dw^A \exp \int d^3x \left[-\frac{g_3^2}{2}\left(\partial_i w^A \partial_i w^A + \lambda w^A w^A\right) - \frac{3\lambda g_3^2}{R}\right].
\end{aligned}
\tag{27}
$$

8.1 Spontaneous mass generation

There is a well-known mechanism of spontaneous mass generation in the $O(n)$ σ-model through the Lagrange multiplier λ getting a nonzero v.e.v. Strictly speaking, the mechanism

is justified at large n but it is plausible that $n = 4$ is large enough for this mechanism to work.

If it does, the four gauge-invariant fields w^A obtain the mass $m_w^2 = <\lambda> \sim g_3^4$. This mass would be observable had we made lattice simulations in the dual formulation (10). However, lattice simulations have been so far done in the usual formulation implying that correlators of only quadratic, quartic,... combinations of the elementary w fields could be measured. Therefore, the observed glueballs are bound states of the w fields. From the quantum numbers, one can infer that the lightest glueballs' masses are

$$m_{0^{++}}, \ m_{2^{++}} \ = \ 2m_w \mp O\left(\frac{1}{n}\right), \tag{28}$$

$$m_{0^{-+}}, \ m_{1^{-+}} \ = \ 4m_w \mp O\left(\frac{1}{n}\right), \qquad n = 4, \tag{29}$$

since the $0, 2^{++}$ operators can be constructed from two w fields while the $0, 1^{-+}$ ones from minimally four. The correction is due to the interaction suppressed as $1/n$. This seems to be in qualitative agreement with the arrangement of the lightest glueballs in $d = 3$. Indeed, according to Teper [20], $m_{0^{++}} \approx 4.7\sqrt{\sigma}$, $m_{2^{++}} \approx 7.8\sqrt{\sigma}$, $m_{0^{-+}} \approx 10\sqrt{\sigma}$, $m_{0^{++}} \approx 11\sqrt{\sigma}$. Of course, one needs to compute the $1/n$ corrections before drawing conclusions.

8.2 Area law

The variables $w^A(x)$ have the meaning of dual gluons (but gauge-invariant!). Mass generation for dual gluons usually implies confinement. Indeed, the average of a large Wilson loop is obtained by adding the mass term into eq. (19):

$$
\begin{aligned}
<W> \ &= \ \int Dw^A \exp\left\{-\frac{g_3^2}{2}\int d^3x \left[(\partial_i w^A + B_i)^2 + m_w^2 \, w^A w^A\right]\right\} \\
&= \ \exp\left\{\frac{g_3^2}{2}\int d^3x \, B_i \left(\delta_{ik}\,\partial^2 - \frac{\partial_i \partial_k}{-\partial^2 + m_w^2}\right) B_k\right\}.
\end{aligned}
\tag{30}
$$

Recalling that $\text{Curl}\,\mathbf{B} = \mathbf{j}$, (see Eqs.(17, 18)) and taking the magnetic field created by a flat circle loop of radius r to be $B_z = \theta(r - \rho)\delta(z)$, $B_{x,y} = 0$, we obtain

$$
\begin{aligned}
<W> \ &\stackrel{r\to\infty}{\Longrightarrow} \ \exp(-\sigma\,\text{Area}), \\
\sigma \ &= \ \frac{3}{32}\,g_3^2 \, m_w.
\end{aligned}
\tag{31}
$$

Estimating m_w from Eq. (28) as half of the average of 0^{++} and 2^{++} masses we get the string tension $\sigma \approx 0.3\,g_3^2$ which is not so far from the lattice value $0.335\,g_3^2$ [20]. Given our negligence of the expected $1/n = 25\%$ corrections to all these formulae, the results for the glueball masses and the string tension are not so bad.

9 Conclusions

1. One can rewrite the quantum YM theory exactly in terms of gauge-invariant variables. In $3d$ $SU(2)$ these are the six coordinates $w^{1-6}(x)$ describing the embedding of the curved dual space into flat space. The theory becomes Einstein's quantum gravity, plus the 'æther' term.

2. Perturbation theory corresponds to expanding about flat dual space. In the lowest order one recovers the Coulomb interaction, plus transverse gluons.

3. The probable infrared regime of the $SU(2)$ theory is that the dual space becomes a S^3 sphere described at low momenta by the $O(4)\,\sigma$-model. In turn, it seems to generate spontaneously the mass for dual gluons, which gives rise to glueball masses and the area law.

4. In $4d$ and/or for $SU(N)$ gauge groups the gauge-invariant variables are also related to the metric of the dual space. The larger N, the higher spin fields are needed to accommodate the necessary number of degrees of freedom of the original YM theory. The action consists of two terms: one has a new type of local symmetry mixing fields with different spins, the other breaks this symmetry but is simple. It is tempting to use this formalism as a starting point for deriving a string from a local field theory.

Acknowledgments

References

[1] A. M. Polyakov, *Nucl. Phys.* **B164**, 171 (1980).

[2] Yu. M. Makeenko and A. A. Migdal, *Phys. Lett.* **B88**, 135 (1979).

[3] A. M. Polyakov, *Yad. Fiz.* **64**, 594 (2001) [*Phys. Atom. Nucl.* **64**, 540 (2001)], hep-th/0006132.

[4] S. Deser and C. Teitelboim, *Phys. Rev.* **D13**, 1592 (1976).

[5] M. B. Halpern, *Phys. Rev.* **D16**, 1798 (1977).

[6] D. Birmingham, M. Blau, M. Rakowski and G. Thompson, *Phys. Rep.* **209**, 129 (1991); J. Baez, e-print gr-qc/9905087.

[7] O. Ganor and J. Sonnenschein, *Int. J. Mod. Phys.* **A11**, 5701 (1996).

[8] D. Diakonov and V. Petrov, *Grav. Cosmol.* **8**, 33 (2002), hep-th/0108097.

[9] F.A. Lunev, *Phys. Lett.* **B295**, 99 (1992).

[10] R. Anishetty, P. Majumdar and H.S. Sharatachandra, *Phys. Lett.* **B478**, 373 (2000), hep-th/9806148.

[11] D. Diakonov and V. Petrov, *Zh. Eksp. Teor. Fiz.* **91**, 1012 (2000) [*J. Exp. Theor. Phys.* **91**, 873 (2000)], hep-th/9912268.

[12] D. Diakonov and V. Petrov, *Phys. Lett.* **B493**, 169 (2000), hep-th/0009007.

[13] E. Witten, *Nucl. Phys.* **B311**, 46 (1988/89).

[14] D. Diakonov and V. Petrov, *J. Exp. Theor. Phys.* **92**, 905 (2001), hep-th/0008035.

[15] D. Diakonov and V. Petrov, *Assimilating the dual space*, invited talk at Channel meeting (Plymouth, Aug. 2002), in preparation.

[16] R. Anishetty et al., *Phys. Lett.* **B 314**, 387 (1993).

[17] I. G. Halliday and P. Suranyi, *Phys. Lett.* **B350**, 189 (1995).

[18] R. Oeckl and H. Pfeiffer, *Nucl. Phys.* **B598**, 400 (2001), hep-th/0008095.

[19] D. Friedan *Ann. Phys. (N.Y.)* **163**, 318 (1985).

[20] M. Teper, *Phys. Rev.* **D59**, 014512 (1999), hep-lat/9804008.