Rigorous results on the ultraviolet limit of non-Abelian gauge theories

Jacques Magnen

Vincent Rivasseau

Roland Sénéor

Centre de physique théorique, CNRS, UPR14
Ecole Polytechnique, 91128 Palaiseau Cedex, France

ABSTRACT

We report on a rigorous construction of the Schwinger functions of the pure SU(2) Yang-Mills field theory in four dimensions (in the trivial topological sector) with a fixed infrared cutoff but no ultraviolet cutoff, in a regularized axial gauge. We review briefly the difficulties related to Gribov ambiguities, then we give an outline of our construction.

A 147 01 92 February 1992

I. Introduction and outline

Non-Abelian gauge theories are used in high energy physics because they are the only field theories renormalizable and asymptotically free. Most physicists are convinced that the ultraviolet problem for these theories is well understood. However it is surprisingly hard to substantiate this belief rigorously beyond perturbation theory, although rigorous constructions of simpler asymptotically free theories such as the Mitter-Weisz or massive Gross-Neveu model in two dimensions have been available for many years now [FMRS1][GK].

Until now there was only one rigorous program completed on this problem, the one of Balaban [B]. This program defines a sequence of block-spin transformations for the Yang-Mills theory in a finite volume on the lattice and shows that as the lattice spacing tends to 0 and these transformations are iterated many times, the resulting effective action on the unit lattice remains bounded. ¿From this result the existence of an ultraviolet limit for *gauge invariant* observables such as "smoothed Wilson loops" should follow, at least for subsequences. Although very impressive, Balaban's work is not easily accessible. Also it does not construct the expectations values of products of the field operators in a particular gauge (the Schwinger functions), because these are not gauge invariant observables. Although physical quantities should be gauge invariant, the Schwinger functions are obviously very convenient for perturbative computations.

A related program in progress but not yet completed is the one of Federbush [F].

In this paper we want to report on our approach to the same problem. We have constructed the Schwinger functions of the pure $SU(2)$ Yang-Mills field in a regularized axial gauge, with a finite volume box as infrared cutoff [MRS]. Our construction remains unfortunately quite complicated and technical. Therefore in this letter we give an outline of the method and results. Our conclusion, after many years of efforts, is the same as Balaban's, namely that the ultraviolet limit of Yang-Mills can be done constructively, although not easily.

Our approach for the moment is limited to the axial gauge. Feynman gauges or similar ones which are Euclidean invariant and convenient for perturbative computations cannot be used directly because of their lack of positivity; this point is discussed in detail below. Also for technical reasons we have to require stability of our ultraviolet cutoff, which rules out certain types of cutoffs. As a justification that our construction is correct we show that the Schwinger functions that we build obey the correct Slavnov identities (with corrections localized at the boundary of our finite volume, since this finite volume cutoff breaks gauge invariance).

We cannot lift the infrared volume cutoff, since this would lead to large values of the coupling constant, hence to non-perturbative effects corresponding to confinement. These problems are for the moment still out of the realm of our methods.

Concerning the axioms of quantum field theory, we cannot study the complete set of Osterwalder-Schrader's axioms, mainly because we never lift our fixed infrared cutoff. However we think that the main axiom, the O.S. positivity, could be shown to hold with some additional work.

Also we do not investigate invariance under large gauge transformations and non-trivial topological effects such as instantons, which are much more difficult to control rigorously from the point of view of mathematical physics.

B) The model. Stabilizing ultraviolet cutoffs.

We consider the pure Euclidean Yang-Mills theory with an infrared cutoff, which we never try to lift. This cutoff may be imposed on the propagator, but we prefer to consider the theory on a finite hypercube Λ with some boundary conditions, which might be periodic, in which case we recover the torus. Such a naive infrared regularization breaks gauge invariance in an explicit way; it creates for instance terms attached to the boundary of Λ in Slavnov identities, and we shall not try to cancel these terms.

Our construction is limited for simplicity to the pure SU(2) Yang-Mills theory, but in principle we think that with some additional work it should apply to any asymptotically free non-Abelian gauge theories (with matter fields). For the pure SU(2) Yang-Mills theory the vector potential is a field $A^a_\mu, \mu = 1, ..., 4, a = 1, 2, 3$ with Lorentz (greek) indices and Lie algebra (latin) indices. Our conventions are those of [IZ]. To simplify the notations we will forget indices most of the time. We write $A = \sum_{a=1}^3 A^a t_a$, with $t_a = (i\sigma_a/2)$ where the σ's are the three usual hermitian Pauli matrices. With this convention the covariant derivative is

$D_\mu = \partial_\mu - g[A_\mu, .]$. We have $\mathrm{Tr}\, t_a t_b = -\frac{\delta_{ab}}{2}$. The field curvature is:

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu) - g[A_\mu, A_\nu] = (\partial \wedge A - g[A, A]) \tag{B.1}$$

g being the coupling constant. The pure Yang-Mills action is (for Euclidean canonical metric $F_{\mu\nu} = F^{\mu\nu}$):

$$-\frac{1}{2}\int_\Lambda d^4x Tr F_{\mu\nu}F^{\mu\nu} = \frac{1}{4}\int_\Lambda d^4x \sum_a F^a_{\mu\nu}F^{\mu\nu a} \tag{B.2}$$

To simplify, we define a scalar product $< A, B >$ by the convention that a trace is taken over all correspondent space time indices and minus a trace over group indices, so that it is positive definite with a factor $1/2$ in component notation. We also write simply A^2 for $< A, A >$, and with this convention we can write the action as $\frac{1}{2}\int_\Lambda F^2$. We distinguish between the quadratic, trilinear and quartic pieces of F^2, introducing a coupling constant g:

$$F^2 = F_2 + gF_3 + g^2F_4 \tag{B.3}$$

This action is invariant under the gauge transformations:

$$A \to A^{\mathbf{g}}\ ;\ (A^{\mathbf{g}})_\mu = \mathbf{g}A_\mu\mathbf{g}^{-1} + g^{-1}\partial_\mu\mathbf{g}\cdot\mathbf{g}^{-1} \tag{B.4}$$

The linearization of (B.4) is:

$$A \to A\ ;\ (A)_\mu = A_\mu + D_\mu \tag{B.5}$$

where $D = \partial - g[A, .]$ is the covariant derivative and takes values in the Lie algebra.

As is usual in constructive theory we try to define the limit of the Schwinger functions with cutoff. They are the moments of a regularized functional measure. This regularized measure has to be well-defined and to coincide in some formal way if the cutoffs are removed with the formal measure of the Yang-Mills theory:

$$e^{-(1/2)\int F^2} \prod_{x\in,a,\mu} DA^a_\mu(x) \tag{B.6}$$

where the infinite product of Lebesgue measures in (B.6) is ill defined.

Due to the invariance (B.4) there is a big number of flat directions in this formal measure. If we want to define only gauge invariant observables, like in Balaban's program one can hope that they will divide out in the normalization. But we want to build Schwinger functions which are not gauge invariant, hence we have to fix the gauge. The nicest choice would certainly be a gauge directly suited for perturbation theory, in which (B.6) becomes

$$e^{-(1/2)\left(\int F^2 - (\partial_\mu A_\mu)^2\right) + \partial_\mu\bar{}\partial_\mu + g\partial_\mu\bar{}[A_\mu,]} \prod_{x\in,a} d^a(x)d^a(x) \prod_{x\in,a,\mu} \bar{d}A^a_\mu(x) \text{(B.7)}$$

where ¯and are the usual Fadeev-Popov ghosts. For $= 1$ we have the Feynman gauge, for $= +\infty$ the Landau gauge; in the case of $SU(2)$ a particularly convenient choice is to fit to a value close to $3/13$ at which the wave function renormalization cancels exactly; we call this gauge the homothetic gauge.

The first task is to build a well defined analogue of (B.7), by introducing some ultraviolet cutoff. The lattice regularization is famous and has many advantages; it is easy to show that the functional integrals which take place in the compact SU(2) group are well defined. Also high temperature expansions show interesting phenomena such as Wilson's area law etc... Since the group variable is $e^{g^{-1}A}$, we can consider that the lattice regularization imposes both an ultraviolet cutoff and an effective limit of order $O(g^{-1})$ on the size of A. In this sense this cutoff is stabilizing at large field. As explained below, a stabilizing cutoff is necessary but not sufficient. One has to understand how the stability is preserved by the multiscale analysis. For instance in the lattice case one has to use block spin transformations which preserve the fact that the rescaled variables at lower momentum remain on a compact group and in Balaban's works, this is the source of many technical complications.

In order to enlarge the possible cutoffs and also to clarify the last point, we have searched for other stabilizing cutoffs in which the stability comes from some explicit operator which is renormalized in the ordinary way in the multiscale analysis. Since Gaussian measures for positive quadratic forms are well defined (e.g. through Minlos's theorem) the standard method in constructive theory is to absorb the propagators of the fields in such measures. On these propagators there are many ways to implement momentum ultraviolet cutoffs (such as Pauli-Villars etc...). Then one has to show that the remaining interaction is stable for large fields. Here this is already a difficult task, because the $F_3 + F_4$ term in (B.3) is not positive by itself when separated from F_2. But it is well known that momentum cutoffs break gauge invariance. We make from this defect a virtue and remark that this gauge breaking can be cancelled by the insertion of appropriate gauge variant counterterms in (B.7). It is only necessary to insert the counterterms corresponding to relevant and marginal operators, since the others have no effects on the theory at finite scale when the cutoff is removed. Using global Euclidean and gauge invariance the important counterterms to consider in the case of SU(2) reduce to A^2 and A^4. Furthermore it is enough to compute the A^4 counterterm at one loop (further contributions being by asymptotic freedom "logarithmically irrelevant"). We performed this task in collaboration with J. Feldman for ultraviolet cutoffs of compact support in momentum space and found that for many shapes of the support function (for instance the shape show in Fig.1 with K sufficiently large) the A^4 counterterm has the right sign so that the theory is stabilized at large field. The detailed analysis can be found in [S1], [R] or [MRS]. It means that the corresponding functional integral with cutoff and gauge restoring counterterms is well defined (recall that F_4 is positive, that F_3 although not positive is only cubic, and that functional integrals for fermions with cutoffs can be explicitly defined through their perturbative expansion, see e.g. [FMRS1][R]). Since the A^4 counterterm is at one loop in $g^4 A^4$ it provides obviously some cutoff on A at size g^{-1}, hence it is in this respect very similar to the lattice cutoff.

Fig.1: The ultraviolet cutoff in momentum space

Using such a cutoff at a scale M where M is some integer we have a reasonable bare ansatz for the theory:

$$e^{-(1/2)\left(\int F_3+F_4+CT(A)\right)+\partial_\mu \bar{\partial}_\mu + g\partial_\mu \bar{[}A_\mu ,]\right)} d\mu(A, \bar{)} \tag{B.8}$$

where $d\mu$ is a certain Gaussian measure which incorporates the propagators of the gauge fixed theory with cutoffs, and CT are the gauge restoring counterterms which at large A behave as $+c.g^4 A^4$. In (B.8) we can take the bare coupling as

$$g_\rho = \frac{1}{\beta_2 (Log M)\rho + \beta_3/\beta_2 \log \rho + C} \tag{B.9}$$

where C is a large constant, and β_2 and β_3 are the usual first non vanishing coefficients of the β function. Then one expects to land on a renormalized coupling constant g_{ren} (the last one in a sequence of effective constants) finite and small if C is large.

C) The Gribov problem

Once a stabilizing cutoff has been introduced the bare functional integral is at least well defined. Unfortunately we have not been able to remove thew cutoff for an ansatz such as (B.8), namely a bare theory with a stabilizing cutoff in a perturbative gauge such as the Feynman or homothetic gauge. The problem we met under the form of a lack of positivity is related to the Gribov phenomenon.

To perform the ultraviolet limit in (B.8) we have to perform on (B.9) a multiscale analysis in the style of [R] and it is crucial that the renormalization group flow of g is dominated by the regular perturbative terms. For this we need to prove that the interaction terms in (B.9) at least in a given momentum slice are small when compared to the Gaussian measure. Equivalently we need a non perturbative bound in the regions $A \simeq g^{-1}$ which tells us that the functional weight of these regions is small compared to the (Gaussian) weight of the region $A \simeq 0$. Although the A^4 term shows that the corresponding functional weight is bounded, we cannot prove solely with this term that this weight is small. This problem is related to the Gribov phenomenon.

Gribov discovered [G] that in the Landau gauge there can be different smooth field configurations which are nevertheless related by a gauge transformation. A configuration $A_2 \neq A_1$ such that $\partial_\mu A_{\mu 1} = \partial_\mu A_{\mu 2} = 0$ and such that there exists a gauge transformation g with $A_1 = A_2^{\mathbf{g}}$ ($A^{\mathbf{g}}$ being defined as in (B.4)) is called a Gribov copy of A_1. There are always some configurations which have copies; this is true even on a compact space and for configurations with the same topological properties, hence inside a given topological sector; it is also true for any regular gauge [S2], not just the Landau gauge. What we call the strong Gribov phenomenon is when there are Gribov copies of the 0 configuration, hence pure gauges which satisfy the gauge condition. This is possible in an infinite space when a change of topological sector or weak decay at infinity is allowed, as shown explicitly in [G], but this strong Gribov phenomenon does not occur inside a given topological sector in a compact space, or under strong decay conditions at infinity if the space is not compact (see e.g. [R]).

Explicit proofs of absence of the strong Gribov phenomenon for finite volume show that there is some definite positivity which lies in the combination of the action F^2 and the Feynman gauge term $(\partial_\mu A_\mu)^2$. Unfortunately this positivity is too weak when the frequencies considered are much larger than the inverse size of the volume cutoff. In intuitive terms, this positivity is tied to boundary conditions: it is useful for the last (physical) momentum slices in a phase space analysis but is not strong enough to help at high momenta.

Although in the constructive study of the ultraviolet limit of non-Abelian gauge theories we can avoid the strong Gribov phenomenon just as we avoid the infrared problem (i.e. by appropriate boundary conditions or compactification), we cannot avoid the existence of copies, in particular in the vicinity of null vectors of the Fadeev-Popov operator

$$K = -\partial_\mu D_\mu = -\Delta + g\partial_\mu [A_\mu, .] \tag{C.1}$$

For $g = 0$ the Fadeev Popov operator reduces to minus the Laplacian and is positive definite once the 0-momentum mode (hence translation invariance) has been deleted. But for non zero g and A it is possible to show that rescaling $A \to k \cdot A$ one can always have negative eigenvalues of K for k large enough (which correspond physically to bound states of the ghost system). The configurations where $\det K = 0$, hence where there exists null vectors for K are the so-called Gribov horizons. The regions inside the first Gribov horizon, where K is positive is called the first Gribov region and so on. In [G] it is shown that near a Gribov horizon there are typically Gribov copies, one on each side of the horizon. These copies can be rapidly decreasing at infinity (or smooth on a compact space) so this "weak Gribov phenomenon" problem cannot be eliminated by an infrared cutoff or by topological restrictions.

The existence of copies means that the functional measure is not monotonous. The fact that the Fadeev-Popov operator is not always positive definite at large fields is also quite disturbing. Since the determinant of this operator occurs in the functional measure, we must conclude that the ordinary formula for functional integration is a signed measure.

It is argued in [H] that this signed measure, although derived in an incorrect way, is nevertheless the correct one; essentially the argument is that as we move away from the gauge orbit of the origin (which, by absence of the strong Gribov phenomenon in our case, cuts the gauge condition only once) the Gribov copies, by some continuity argument, should occur in pairs with equal and opposite values, therefore cancel out so that the naive prescription with the signed determinant is equivalent to the correct prescription in which a single point on each gauge orbit is selected. In any case even if this conjecture is true it seems difficult to prove it in constructive theory.

An other possibility is to eliminate Gribov copies at the beginning by use of a better ansatz. This is the approach advocated e.g. by Zwanziger [Z], in which one tries to define first a correct configuration space for the functional measure, eliminating Gribov copies by additional gauge conditions. It seems very reasonable to consider that the functional integration for the non-Abelian gauge theory can be written as a positive measure supported on a subdomain of the first "Gribov region". However to characterize this region in a constructive way seems hard. Therefore we have turned to the axial gauge which contains some definite positivity and lacks the Gribov effects.

D) The regularized axial gauge

The axial gauge is defined by the condition $A_0 = 0$. This gauge condition (which is not a complete gauge fixing) can be implemented without paying the price of a Fadeev-Popov determinant (more precisely this determinant is a constant which disappears in the normalization). In this gauge the action is

$$e^{-(1/2)(<A,p_0^2 A>+F_{sp}^2)} \tag{D.1}$$

where F_{sp} is the sum over spatial indices (excluding the time component).

It is easy to verify that when coupled to a stabilizing cutoff which states that A is of order g^{-1} at most, the positivity in (D.1) implies that A is of order $g^{-1/2}$ in probability. Indeed if we use $g^2 < A, p^2 A >$ as stabilizing term (it is marginal, like A^4, and weaker at large A since it is only quadratic instead of quartic) we can join it to $e^{-(1/2)<A,p_0^2 A>}$ and obtain a covariance of order g^{-1}.

For $A \simeq g^{-1/2}$ we are well within the first Gribov region, and F_2 is much larger than F_3 and F_4. This means that both perturbation theory and an explicit change of gauge become possible. However the use of the positivity contained in the $< A, p_0^2 A >$ term has again a technical price attached to it: we must perform the phase space analysis of the theory in an anisotropic way, with momentum slices which distinguish both on the value of p^2 and p_0^2. The corresponding dual cells in x-space are rectangular, elongated in the time direction, and this is the source of many complications.

E) The expansion and the normalization of large field regions

A phase space expansion consists in both spatial cluster expansions and momentum decoupling expansions. The only important restriction to build these interpolations is to preserve positivity requirements for the interpolated theory.

In addition to perturbative computations, constructive theory needs some stability properties of the functional integral at large fields. When a multiscale analysis is necessary, which is certainly the case in a renormalizable theory, we need to decouple functional integrals at different scales and this results in a typical problem called domination. The low momentum fields produced by the expansion, if evaluated with the Gaussian measure, would lead to the divergence of perturbation theory and we must instead bound them by using the positivity of the effective potential. In our case we cannot use the perturbative small field effective potential when the low momentum fields are large. In other words there are some couplings which are not dominable. Fortunately the structure of these non dominable terms is simple. They can be absorbed simply by changing ordinary derivatives of the field into covariant derivatives corresponding to the background field which is the sum of all large field of lower momenta. This rule is a deep consequence of the geometric structure of the theory.

Applying this idea we have to replace derivatives by covariant derivatives D even in the gauge fixing terms. This reacts on the Fadeev-Popov determinant. Also the multiscale analysis as to be performed really around the zeroes of the operator D^2 rather than the ordinary Laplacian.

Furthermore we want the ultraviolet cutoff to be stabilizing in the sense above. The counterterms being computed by a one loop analysis which is easily performed for a perturbative gauge like (B.7), we choose to use a simple form of the cutoff after gauge transformation to the perturbative gauge. This means that the true cutoff in the axial gauge is given by an inverse formula applied to a simple momentum cutoff. These two effects unfortunately complicate very much the starting ansatz for our axial gauge theory, and we refer to [MRS] for the corresponding formulas.

Then the multiscale analysis of the theory proceeds in the usual way except for the following complications: there is a division of the functional integral cells corresponding to the phase space analysis into large field cells and small field cells. The large field cells $A >> g^{-1/2}$ are suppressed in probability by the positivity of the axial gauge. In the small field cells we perform an explicit change of gauge by a well-defined analogue of the Fadeev-Popov formula with cutoff, and we obtain a theory with a perturbative gauge analogous to the homothetic gauge ($\simeq 3/13$ in (B.7)), but with covariant derivatives in the background low-momentum large field instead of ordinary derivatives. The propagator is therefore not translation invariant, and to perform the usual cluster expansion analysis and renormalization in each momentum slice requires an auxiliary expansion, in which one compares the propagator with a slowly varying background field to the propagator in a constant background field. But we have also a last problem, typical of a large versus small field expansion. To be negligible in the correct sense the functional weight of the large field cells has to be not only small, but so small that it compensates for the difference in the normalization of these cells. The small field cells have indeed a normalization corresponding to their background dependent Gaussian measure, whereas the large field cells have the normalization of the initial axial gauge.

In fact for every large field cell there is an associated domain of small field cells for which this large field B is a part of the background. The difference in normalization of this small field domain with and without the background field B gives rise to a determinant which can and in fact must be associated to the large field cell (more precisely it gives a quotient of determinants if the Fadeev-Popov terms is also taken into account). The crucial point to check is that the large field condition together with the positivity of the axial gauge gives a factor so small that it compensates for the (non-perturbative) bound that we have on this normalizing ratio of determinants. The fact that this ratio of determinants can be explicitly bounded is related to the fact that its leading behavior when the small field domain contains many scales is again given by the relevant term in B^2, whose coefficient in the determinant can be explicitly computed; we found that in the case of a stabilizing ultraviolet cutoff the corresponding bound can be controlled.

In conclusion the corresponding expansion converges if the renormalized coupling at the infrared scale remains small (e.g. when C in (B.9) is big enough), and the Schwinger functions in the limit satisfy the (non-perturbative) Slavnov identities of the axial gauge, with a right hand side depending on the boundary conditions on the volume cutoff, which represents the infrared gauge breaking effects. The gauge transformed functions corresponding to the small field regions hence to the perturbative gauge also satisfy the corresponding Slavnov identities at every order in perturbation theory, but not non-perturbatively, since the large field region give exponentially small correction terms.

Acknowledgements We thank warmly J. Feldman for his collaboration at an early stage of this work.

REFERENCES

[B] T. Balaban, Comm. Math. Phys. 95, 17 (1984), 96, 223 (1984), 98, 17 (1985), 99, 75 (1985), 99, 389 (1985), 102, 277 (1985), 109, 249 (1987), 116, 1 (1988), 119, 243 (1988), 122, 175 (1989), 122, 355 (1989).

[F] P. Federbush, A phase cell approach to Yang-Mills theory I-VII, Comm. Math. Phys 107, 319 (1986), 110, 293 (1987), 114, 317 (1988); Ann Inst. Henri Poincaré, 47, 17 (1987).

[FMRS1] J. Feldman, J. Magnen, V. Rivasseau and R. Sénéor, A renormalizable field theory: the massive Gross-Neveu model in two dimensions, Comm. Math. Phys. 103, 67 (1986).

[G] V. Gribov, Quantization of non-Abelian gauge theories, Nucl. Phys. B139, 1 (1978).

[GK] K. Gawedzki and A. Kupiainen, Gross-Neveu model through convergent perturbation expansions, Comm. Math. Phys. 102, 1 (1985).

[H] P. Hirschfeld, Strong evidence that Gribov copying does not affect the gauge theory functional integral, Nucl. Phys. B157, 37 (1979).

[IZ] C. Itzykson and J.B. Zuber, Quantum field theory, Mc Graw and Hill, New-York, 1980.

[MRS1] J. Magnen, V. Rivasseau and R. Sénéor, Construction of YM_4 with an infrared cutoff, preprint Ecole Polytechnique, 1992.

[R] V. Rivasseau, From perturbative to constructive renormalization, Princeton University Press, 1991.

[S1] R. Sénéor, in Renormalization of Quantum Field Theories with Non-Linear Field Transformations, P. Breitenlohner, D. Maison and K. Sibold Eds, Lecture Notes in Physics n^o 303, Springer Verlag, 1988.

[S2] I. Singer, Some remarks on the Gribov Ambiguity, Comm. Math. Phys. 60, 7 (1978).

[Z] D. Zwanziger, Non-perturbative modification of the Fadeev-Popov formula and banishment of the naive vacuum, Nucl. Phys. B209, 336 (1982), Action from the Gribov horizon, Nucl. Phys. B321, 591 (1989), Local and renormalizable action from the Gribov horizon, Nucl. Phys. B323, 513 (1989).