5
Yang–Mills theories

5.1 Introduction

Since the unification of the electromagnetic and weak interactions through the Glashow–Salam–Weinberg model [75], Yang–Mills theories [76] have been widely accepted as correctly describing elementary particle physics. This belief was reinforced when they proved to be renormalizable [77, 78]. Moreover, the discovery of color symmetry as the underlying gauge invariance associated with strong interactions raised the possibility that all interactions of nature could possibly be cast as Yang–Mills theories. This spawned interest in grand unified models and some partial successes were achieved in this direction.

A crucial ingredient in the description of elementary particle physics through gauge theories is the maintenance of the gauge invariance of physical results and the underlying theory and this is also crucial in order to be able to prove renormalizability.

The success of the electroweak model is yet to be achieved by the quark model of strong interactions. The reason is that perturbative techniques, which were adequate for the electroweak model, are only appropriate in the high energy regime of strong interactions. This motivated the interest in non-perturbative techniques, especially to prove the existence of a confining phase. A great effort took place in the late 1970s and suggestive arguments were put forward but a rigorous proof of quark confinement is still lacking.

In several of these attempts the use of loops played an important role. Loops were used in a variety of contexts and approaches including the one we are focusing on in this book, the loop representation. In this chapter we will also briefly highlight some of the aspects of other approaches which seem of most interest for gravitational physicists. We are forced to omit, for reasons of space, many other valuable constructions.

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

The first gauge invariant, path dependent formulation of a gauge theory was Mandelstam's reformulation of QED [8]. Mandelstam later extended his formulation to the Yang–Mills case and applied it to the development of Feynman diagrammatic rules [9]. This was the first time the Feynman rules for non-Abelian gauge theories had been found through canonical quantization. They had been established in the S-matrix approach by Feynman [79] and DeWitt [80] and in the functional approach by Fadeev and Popov [81]. The main feature of the Mandelstam approach was to avoid using gauge dependent quantities, introducing instead path dependent field variables $F_{ab}(P)$ for the field where P is a path going from a basepoint to the point of interest: translating to the language introduced in chapter 1

$$F_{ab}(\pi) = \lim_{\gamma\to\iota} \Delta_{ab}(\pi_o^x) H_\gamma[A]. \tag{5.1}$$

These quantities satisfy the identities induced by those of the loop derivative that we introduced in chapter 1 and the Yang–Mills equation of motion, $D_a F^{ab}(P) = 0$, where D^a is the Mandelstam covariant derivative. Notice that the aim of this approach was to develop a perturbative formulation and in that respect it was successful.

Another approach was that of Polyakov [83, 82]. This was based on the hope that holonomies for Yang–Mills theories could satisfy equations similar to those of non-linear σ models, which, in turn, are integrable. This is based on what happens in $2+1$ dimensions. The basic variable is a derivative of the holonomy

$$F_\mu(s,\gamma) = \frac{\delta H_\gamma[A]}{\delta \gamma^\mu(s)} H_\gamma^{-1}[A] \tag{5.2}$$

and the formalism assumes a parametrization has been picked for the loop and extra equations are added to impose invariance under reparametrizations. The equations of motion are

$$\frac{\delta F_\mu(s,\gamma)}{\delta\gamma^\nu(t)} - \frac{\delta F_\mu(t,\gamma)}{\delta\gamma^\nu(s)} + [F_\mu(s,\gamma), F_\nu(t,\gamma)] = 0, \tag{5.3}$$

$$\frac{d\gamma^\mu(s)}{ds} F_\mu(s,\gamma) = 0, \tag{5.4}$$

$$\frac{\delta F_\mu(s,\gamma)}{\delta\gamma^\mu(s)} = 0. \tag{5.5}$$

The first equation is the usual vanishing of a curvature that appears in non-linear σ models. The second equation is related with the invariance under reparametrizations of the holonomy and the last equation is a consequence of the Yang–Mills dynamical equation.

This approach had several difficulties. Even in the three-dimensional case, the equations presented are not exactly the same as those of a traditional non-linear σ model. In a traditional non-linear σ model, the first equation would involve a partial derivative with respect to a coordinate. In the present case, this means that one is really dealing with an infinite number of components, one per each point in parameter space, as is expected in loop space. The third equation, which in the usual case is a divergence, should be summed over all components (integrated over s), but then, it is not true that the Yang–Mills equations follow. This difficulty was recognized by Polyakov [84]. Moreover the situation becomes more complicated if one considers the four-dimensional case, since in that case it is not even clear how to reformulate the fields as σ models. Other technical difficulties appear, mainly related to the parametrization dependence [85]. In particular it was shown that when the equations are rewritten in a parametrization independent way (using the techniques discussed in chapter 1) extra terms appear, which break the resemblance with the σ model.

The plan of this chapter is as follows. We will discuss in some detail in the next section an alternative approach, due to Polyakov and Migdal. We then devote a section to the loop representation of Yang–Mills theories, discussing the $SU(2)$ and $SU(N)$ cases. We end with a section on some ideas relating loops to confinement.

5.2 Equations for the loop average in QCD

The approach we are about to discuss originated in an idea of Polyakov [82] and was later developed by Makeenko and Migdal [12, 86]. We only present a sketch of the main ideas here, in part because we will use similar techniques in the context of Chern–Simons theory in chapter 10. We refer the reader to the review article by Migdal [11].

The basic idea is as follows. The expectation value of the Wilson loop functional in (Euclidean) four dimensions (i.e, the loop exists in a four-dimensional space) operates as a generating functional of the Green functions of the theory, as can be simply seen by considering its successive loop derivatives at different points,

$$\Delta_{\mu_1\nu_1}(\pi_o^{x_1})\ldots\Delta_{\mu_n\nu_n}(\pi_o^{x_n}) < W(\gamma) >_{\gamma=\iota} = < \mathrm{Tr}[F_{\mu_1\nu_1}(x_1)\ldots F_{\mu_n\nu_n}(x_n)] > . \tag{5.6}$$

Notice that to write the right-hand side as point dependent a prescription for the paths π has been chosen (as we discussed at the end of chapter 1).

The right-hand side of expression (5.6) is the n-point function of the theory. This was the insight of Polyakov. Now consider the action of the field equations on the expectation value of the Wilson loop functional,

$$D^\mu\Delta_{\mu\nu}(\gamma_o^x) < W(\gamma) > = \int DA\, \exp(-S_{YM}) D^\mu\Delta_{\mu\nu} W_A(\gamma)$$

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

$$= \int DA \, \exp(-S_{YM}) \mathrm{Tr}[D^\mu F^i_{\mu\nu}(x) \mathbf{X}^i U_A(\gamma^x_x)]$$
$$= -\int DA \frac{\delta}{\delta A^i_\nu} \exp(-S_{YM}) \mathrm{Tr}[\mathbf{X}^i U_A(\gamma^x_x)]$$
$$= \int DA \, \exp(-S_{YM}) \frac{\delta}{\delta A^i_\nu} \mathrm{Tr}[\mathbf{X}^i U_A(\gamma^x_x)]$$
$$= \int DA \, \exp(-S_{YM}) \oint_\gamma dy^\mu \delta^4(x-y)$$
$$\times \mathrm{Tr}[\mathbf{X}^i U_A(\gamma^y_x) \mathbf{X}^i U_A(\gamma^x_y)] \tag{5.7}$$

where S_{YM} is the Yang–Mills action $\frac{1}{2}\int d^4x \mathrm{Tr}[\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}]$ and $\mathbf{X}^i$ are the generators of the group.

Let us now particularize the gauge group to $SU(N)$. This allows us to use the identity

$$\mathbf{X}^i_{AB} \mathbf{X}^i_{CD} = (\delta_{AD}\delta_{BC} - \tfrac{1}{N}\delta_{AB}\delta_{CD}) \tag{5.8}$$

and to reexpress the above result as

$$D^\mu \Delta_{\mu\nu}(\gamma^x_o) <W(\gamma)> =$$
$$\oint_\gamma dy^\mu \delta^4(x-y)(<W(\gamma^y_x)W(\gamma^x_y)> - \frac{1}{N} <W(\gamma)>). \tag{5.9}$$

Notice that this equation couples the expectation value of the Wilson loop functional with the expectation value of products of Wilson loops. In general one would therefore need to consider similar equations for $< W(\gamma_1) \ldots W(\gamma_n) >$. However, in the particular case of $SU(2)$ or $\lim_{N\to\infty} SU(N)$ it is enough to consider only the expectation value of one Wilson loop functional. In the $SU(2)$ case this is justified since one can reexpress any product of Wilson loop functionals in terms of a single Wilson loop. In the $N \to \infty$ case it can be shown [12] that,

$$< W(\gamma_1) \ldots W(\gamma_n) > = < W(\gamma_1) > \ldots < W(\gamma_n) > + O(N^{-2}), \tag{5.10}$$

due to the fact that the leading Born terms correspond to the sum of all planar diagrams [87]. The Makeenko–Migdal equation can be rewritten for this particular case as

$$D^\mu \Delta_{\mu\nu}(\gamma^x_o)\phi(\gamma) = \oint_\gamma dy^\mu \delta^4(x-y)\phi(\gamma^y_x)\phi(\gamma^x_y), \tag{5.11}$$

where

$$\phi(\gamma) = < W(\gamma) > . \tag{5.12}$$

This equation is reminiscent of that of a $\lambda\phi^3$ scalar field theory. Notice that the equation is only non-trivial if one considers intersecting loops. For smooth loops the right-hand side of the equation becomes $\phi(\gamma)$ and the solution to the equation can be found and coincides with the vacuum state of Maxwell theory in terms of loops that we introduced in the previous chapter. That is, the non-Abelian character of the theory is lost if one does not consider intersecting loops, a fact we will also see reflected in the Hamiltonian case.

It can be shown [88] that the Makeenko–Migdal equations generate all planar diagrams in perturbation theory in a regularized fashion (if one regularizes the equation), although they are not renormalized and no concrete proposal has been found for an equation that could take care of the renormalization.

This approach offered the promise of reformulating QCD entirely in terms of free color fields, which raised the hope that confinement could be understood. Moreover, it makes it possible to express the expectation values of the observables through integrals in loop space. The diagrams are automatically free of infrared catastrophes since one works only with gauge invariant quantities. Finally, Migdal [12] gave a heuristic argument that showed that the behavior of the Wilson loop is consistent with the asymptotic area law typical of confinement.

Several obstacles hampered further development of this approach. To begin with, the expectation values considered are divergent and need to be renormalized, as can be seen from their perturbative study [88]. The equation was initially written [12] in terms of a functional derivative, which led to some technical problems, though a later more geometric reformulation was accomplished [89]. Historically, for a long time the structure and completeness of Mandelstam identities were not understood for the different gauge groups. Although the conceptual simplicity introduced by expressing observables as integrals in loop space was appealing, it is also the case that one does not know how to compute such integrals.

But the main obstacle in this approach is the fact that not a single solution of the Makeenko–Migdal equation has ever been found in four dimensions. Progress has been made in the two-dimensional case [91, 90, 93] and also with variational techniques [92]. It would be interesting to test whether the ideas of the extended representation we present in this book can be used to tackle this problem. Another interesting aspect is that relations have been found between the Makeenko–Migdal equation in the large N limit and equations for string theory [94].

Most of the advantages and disadvantages of Migdal's approach are shared by other loop formulations. The appeal of the loop representation, which is the main theme of this book, lies elsewhere. On the one hand, loop representations based on Hamiltonian approaches deal with three-dimensional loop equations in a realistic case instead of four-dimensional ones as in the Migdal construction. Moreover, they are better suited for a canonical description of quantum gravity. In particular one can solve the diffeomorphism constraint rather easily, as we will see in chapter 8. We will also see in chapter 10 that by applying exactly the same construction that we presented here for the Yang–Mills case to the Chern–Simons action, one can find the connection between the expectation value of a Wilson loop and the Jones polynomial of knot theory.

5.3 The loop representation

Constructing a loop representation for a Yang–Mills theory is a straightforward matter with the concepts introduced in chapter 3.

We will first discuss the $SU(2)$ case as an illustration of the quantization of a non-canonical algebra. We then discuss the $SU(N)$ case via a transform.

5.3.1 $SU(2)$ Yang–Mills theories

Let us construct the quantum theory for $SU(2)$ through the quantization of a non-canonical algebra of loop dependent operators. A quantization could be achieved (formally) using the loop transform, leading to the same quantum theory.

Let us consider a gauge theory with $SU(2)$ gauge group. The connections are group-valued $A_a = A_a^i \mathbf{X}^i$ where the Xs are elements of the $su(2)$ algebra. We define the following quantities

$$T^0(\gamma) = \mathrm{Tr}[U(\gamma)] = \mathrm{Tr}\left[P\exp\left(ig\oint_\gamma dy^a A_a\right)\right], \tag{5.13}$$

$$T^a(\gamma_x^x) = \mathrm{Tr}[U(\gamma_o^x)\tilde{\mathbf{E}}^a(x)U(\gamma_x^o)], \tag{5.14}$$

in the same spirit as those defined in chapter 3, except that we are making explicit the dependence on the coupling constant g of the theory.

Classically, these quantities satisfy an algebra under Poisson brackets, which we discussed in chapter 3,

$$\{T^a(\gamma_x^x), T(\eta)\} = \frac{i}{2}\sum_{\epsilon=-1}^{1} \epsilon X^{ax}(\eta)T(\gamma\circ\eta^\epsilon), \tag{5.15}$$

$$\begin{aligned}\{T^a(\gamma_x^x), T^b(\eta_y^y)\} = &-\frac{i}{2}\sum_{\epsilon=-1}^{1} \epsilon X^{ax}(\eta)T^b(\eta_y^x\circ(\gamma_x^x)^\epsilon\circ\eta_x^y),\\ &+\frac{i}{2}\sum_{\epsilon=-1}^{1} \epsilon X^{by}(\gamma)T^a(\gamma_x^y\circ(\eta_y^y)^\epsilon\circ\gamma_y^x),\end{aligned} \tag{5.16}$$

where η^ϵ represents either η or η^{-1}.

The best picture of these relations can be obtained by considering the $T^a(\gamma_x^x)$ to be represented by a loop with a "hand" at the point x. Then the commutators are only non-zero when the hand on one of the loops grabs the other loop (which implies the loops must intersect). The effect of "grabbing" one loop with the hand of the other is to insert the accompanying loop at the point of intersection.

This algebra is closed, as seen above, but it is insufficient in the sense that one cannot express all observables of interest in Yang–Mills theories in terms of it, in particular, the Hamiltonian. Therefore one cannot base a quantum theory simply on finding a representation of this algebra. One should consider a larger algebra including the objects with n insertions defined by formula (3.102) that were discussed in chapter 3. The algebra of T^0 and T^a has for this reason been called the "small algebra" [38]. A generic Poisson bracket between higher order Ts is given by

$$\{T^{b_1\ldots b_n}(\gamma_{x_1}^{x_2},\ldots,\gamma_{x_n}^{x_1}),T^{a_1\ldots a_m}(\eta_{y_1}^{y_2},\ldots,\eta_{y_m}^{y_1})\}=\frac{i}{2}\sum_{\epsilon=-1}^{1}$$
$$\Bigg\{-\sum_{k=1}^{m}\epsilon X^{b_k\,x_k}(\eta)T^{b_1\ldots \not{b}_k\ldots b_m,a_1\ldots a_n}(\eta_{y_1}^{y_2},\ldots,\eta_{y_{k-1}}^{x_k}\circ(\gamma_{x_k}^{x_k})^{\epsilon}\circ\eta_{x_k}^{y_k},\ldots,\eta_{y_m}^{y_1})$$
$$+\sum_{k=1}^{n}\epsilon X^{a_k\,y_k}(\eta)T^{a_1\ldots \not{a}_k\ldots a_n,b_1\ldots b_m}(\gamma_{x_1}^{x_2},\ldots,\gamma_{x_{k-1}}^{y_k}\circ(\eta_{y_k}^{y_k})^{\epsilon}\circ\gamma_{y_k}^{x_k},\ldots,\gamma_{x_m}^{x_1})\Bigg\},\tag{5.17}$$

where $\not{b}_k$ means that the index is not present.

The Hamiltonian for a Yang–Mills theory was introduced in chapter 3, and is given by

$$\tilde{\tilde{\mathcal{H}}}=\int d^3x\mathrm{Tr}(\tfrac{1}{2}(\tilde{\mathbf{E}}^a\tilde{\mathbf{E}}^b\eta_{ab}+\tilde{\mathbf{B}}^a\tilde{\mathbf{B}}^b\eta_{ab})).\tag{5.18}$$

The two terms in this expression have different properties. The $\mathbf{B}^2$ term can actually be written in terms of a $T^0(\gamma)$ as

$$\mathrm{Tr}(\tilde{\mathbf{B}}^a(x)\tilde{\mathbf{B}}^b(x))\eta_{ab}=-\tfrac{1}{2}\lim_{\eta\to\iota}\eta^{ac}\eta^{bd}\Delta_{ab}(\pi_o^x)\Delta_{cd}(\pi_o^x)T^0(\gamma),\tag{5.19}$$

i.e., by taking a double loop derivative and then shrinking the loop dependence to a point. The right-hand side of this formula does not depend on the path choice, as can be seen explicitly by considering

$$\Delta_{ab}(\pi_o^y)\Delta_{cd}(\pi_o^x)T^0(\gamma)=\mathrm{Tr}[U(\pi_o^y)F_{ab}(y)U(\pi_y^o)U(\pi_o^x)F_{ab}(x)U(\pi_x^o)U(\gamma)]\tag{5.20}$$

and taking the limit in which $x\to y$ and the limit of γ shrinking to a point giving,

$$\Delta_{ab}(\pi_o^x)\Delta_{cd}(\pi_o^x)T^0(\gamma)\mid_{\gamma\to\iota}=Tr[F_{ab}(x)F_{cd}(x)],\tag{5.21}$$

which is independent of the path prescription.

The term involving two electric fields cannot be written in terms of T^0 or T^a, one needs a T^{ab}. Classically this can be seen from

$$\mathrm{Tr}(\tilde{\mathbf{E}}^a(x)\tilde{\mathbf{E}}^b(x))\eta_{ab} = \lim_{\gamma\to\iota} \eta_{ab}T^{ab}(\gamma_x^y, \gamma_y^x) \tag{5.22}$$

and in the limit $\gamma \to \iota$ the points x and y coincide.

We now proceed to propose a quantization of the classical non-canonical algebra. We consider a space of wavefunctions of loops $\Psi(\gamma)$ as discussed in section 3.5.3 and define the action of the operators as

$$\hat{T}^0(\eta)\Psi(\gamma) \equiv \Psi(\gamma \circ \eta) + \Psi(\gamma \circ \eta^{-1}), \tag{5.23}$$

$$\hat{T}^a(\eta_x^x)\Psi(\gamma) \equiv -\tfrac{1}{2}\sum_{\epsilon=-1}^{1} \epsilon \oint_\gamma dy^a \delta(x-y)\Psi(\gamma\circ\eta^\epsilon). \tag{5.24}$$

The action of the T^2 is defined as [39]

$$\hat{T}^{ab}(\eta_x^y, \eta_y^x)\Psi(\gamma) = \tfrac{1}{4}X^{ax}(\gamma)X^{by}(\gamma)[\Psi(\gamma_x^y \circ \bar{\eta}_y^x, \gamma_y^x \circ \bar{\eta}_x^y)$$
$$+\Psi(\gamma_x^y \circ \eta_y^x, \gamma_y^x \circ \eta_x^y) + \Psi(\gamma_x^y \circ \bar{\eta}_y^x \circ \bar{\gamma}_x^y \circ \eta_y^x) + \Psi(\gamma_y^x \circ \bar{\eta}_x^y \circ \bar{\gamma}_y^x \circ \eta_x^y)]. \tag{5.25}$$

This last expression could be rearranged in terms of wavefunctions of a single loop using the Mandelstam identity (3.119).

This representation for the T operators yields a quantum commutator algebra that reproduces, to first order in $\hbar$, the classical Poisson algebra of the T operators that we introduced in chapter 3*.

We are now in a position to give a quantum representation of the Hamiltonian of the theory. We have written the Hamiltonian of the theory (5.18) and it is the sum of two terms, one electric and one magnetic. The electric portion was given as a limit of a T^2, so we can now find the corresponding quantum representation, by taking the limit in equation (5.25) in which we shrink the loop η to a point. We will study the action of this operator at a point in the loop γ where there is an intersection (of arbitrary order). The action on a regular point of γ can be obtained as a limit (or by direct calculation). The result is

$$\lim_{\gamma\to\iota} T^{ab}(\eta_x^y \circ \eta_y^x)\Psi(\gamma) = 2X^{ax}(\gamma)X^{by}(\gamma)(\tfrac{1}{4}\Psi(\gamma_x^y\circ\gamma_y^x) + \tfrac{1}{2}\Psi(\gamma_x^y\circ\bar{\gamma}_x^y)). \tag{5.26}$$

In the case of an intersection, the points x and y lie on any "petal" of the loop and therefore the portions γ_x^y and $\bar{\gamma}_y^x$ refer to the various combinations of petals contained between x and y. For the case of a

* We are taking $\hbar = 1$. The orders of $\hbar$ can be restored by noting that momenta are first order in $\hbar$, for instance, $T^1 = O(\hbar)$, $T^2 = O(\hbar^2)$, etc.

regular point of the loop, one simply takes the loop and shrinks the points x to y and therefore $\gamma_x^y \rightarrow \gamma$ and $\gamma_y^x \rightarrow \iota$. The result then is

$$\lim_{\eta \rightarrow \iota} T^{ab}(\eta_x^y \circ \eta_y^x)\eta_{ab}\Psi(\gamma) = \tfrac{3}{4}X^{ax}(\gamma)X^{by}(\gamma)\Psi(\gamma). \quad (5.27)$$

We also see that smooth loops are eigenstates of the electric part of the Yang–Mills Hamiltonian. This allows us to think of Wilson loops as lines of electric flux.

The complete Hamiltonian for an $SU(2)$ Yang–Mills theory in the loop representation is given by

$$\begin{aligned} \hat{\mathcal{H}}\Psi(\gamma) = \Big[& -\tfrac{1}{4}\int d^3x \eta^{ac}\eta^{bd}\Delta_{ab}(x)\Delta_{cd}(x) \\ & + \tfrac{1}{4}\oint_\gamma dy^a \oint_\gamma dy'^b \eta_{ab}\delta^3(y-y')\Big]\Psi(\gamma) \\ & + \tfrac{1}{2}\oint_\gamma dy^a \oint_\gamma dy'^b \eta_{ab}\delta^3(y-y'))\Psi(\gamma_y^{y'} \circ \bar{\gamma}_{y'}^y). \end{aligned} \quad (5.28)$$

The path dependence of the loop derivatives has been dropped since we showed above that they are prescription independent. Notice that if the loop γ does not have intersections, the last term is equal to the second one and the equation is identical (up to constants) to the one obtained for Maxwell theory (4.72). Therefore it is clear that wavefunctions must have support on intersecting loops if the theory is to capture the full non-Abelian nature of the fields.

As in the case of Maxwell theory, the Hamiltonian is singular and needs to be regularized and renormalized. As we pointed out in the case of Maxwell theory, in principle *all terms* in the Hamiltonian require a regularization. In the case of Maxwell theory we knew how to compute the vacuum and this suggested a suitable regularization of wavefunctions and operators. In the non-Abelian case, unfortunately, we do not know a single solution of the Hamiltonian eigenvalue equation and the issue of regularization and renormalization is largely unexplored. The eigenvalue equation has been extensively studied in the lattice in different approximations, leading to results for the energy density, gluon mass spectrum and other observables which coincide with those obtained with more standard methods. We will return to the lattice treatment in the next chapter.

5.3.2 $SU(N)$ Yang–Mills theories

The loop representation for $SU(N)$ Yang–Mills theories can be built along similar lines to the ones we followed in the previous section for the $SU(2)$ case. We will see that the main difference consists in the fact that one needs to consider wavefunctions of multiloops. The classical "small" algebra of T^0 and T^1 can be readily generalized to the $SU(N)$ case,

$$\{T^0(\gamma), T^0(\eta)\} = 0, \tag{5.29}$$

$$\{T^0(\gamma), T^a(\eta_x^x)\} = ig \oint_\gamma dy^a \delta(y-x) \times \left(T^0(\gamma_o^y \circ \eta_x^x \circ \gamma_y^o) - \tfrac{1}{N} T^0(\gamma) T^0(\eta)\right), \tag{5.30}$$

$$\begin{aligned}\{T^a(\gamma_x^x), T^b(\eta_y^y)\} = &-ig \oint_\eta dz^a \delta(z-x) \\ &\times \left(T^b(\eta_y^z \circ \gamma_x^x \circ \eta_z^y) - \tfrac{1}{N} T^0(\gamma) T^b(\eta_y^y)\right) \\ &+ig \oint_\gamma dz^b \delta(z-y) \\ &\times \left(T^a(\gamma_x^z \circ \eta_y^y \circ \gamma_z^x) - \tfrac{1}{N} T^0(\eta) T^a(\gamma_x^x)\right). \end{aligned} \tag{5.31}$$

Up to now, representations of the large algebra have not been studied for $SU(N)$ with $N > 2$. The approach we will take will be to consider the representation of the elements of the small algebra with the addition of a T^{ab}, which is sufficient to write the Hamiltonian. This is clearly not enough: the Poisson bracket of a T^{ab} with a T^{cd} gives rise to T^{abc}s. A quantization of the full algebra can always be performed such that this Poisson bracket is represented by a correct commutation relation. A direct constructive procedure to obtain such an algebra would be to perform a usual canonical quantization in terms of Es and As and to consider the Ts as derived quantities. The resulting quantum algebra will coincide with Poisson bracket algebra up to factor ordering differences. We should remind the reader that in the quantization process Poisson brackets are replaced by i times $\hbar$ and the factor ordering ambiguities are of order $\hbar^2$ or higher. Since the T algebra for $SU(N)$ has not been explicitly computed up to now, we will proceed with the constructive technique we just outlined. It would be interesting to check explicitly that this technique yields the same result as consideration of the full T algebra.

We will choose the following ordering prescription for the T^1 operator in terms of the canonical operators,

$$\hat{T}^a(\gamma_x^x) = \mathrm{Tr}\left(H(\gamma_x^x)\hat{\mathbf{E}}^a\right). \tag{5.32}$$

The operator algebra of $\hat{T}^1$ and $\hat{T}^0$ reproduces the same classical Poisson algebra described above, where the brackets are replaced by $\hbar$ times the commutators.

We now represent this algebra in terms of loop-based operators and wavefunctions, giving rise to the loop representation. As we mentioned in section 3, we need to consider functions of multiloops, via the transform

$$\Psi(\gamma_1,\ldots,\gamma_n) = \int d\mathbf{A}\, W^*_{\mathbf{A}}(\gamma_1)\cdots W^*_{\mathbf{A}}(\gamma_n)\Psi[\mathbf{A}]. \tag{5.33}$$

It turns out to be convenient to consider a certain combination of products of Wilson loops in the transform which is slightly different than the one presented in equation (5.33). We will consider the following transform:

$$\Psi(\gamma_1,\ldots,\gamma_n) = \int d\mathbf{A}\, M^*_N(\gamma_1,\ldots,\gamma_n)\Psi[\mathbf{A}], \tag{5.34}$$

where the functionals $M_N(\gamma_1,\ldots,\gamma_n)$ were introduced in section 3.4.1. This does not imply any loss of generality since the product of N Wilson loops can be reconstructed from the Ms.

We now define the action of the T^0 operator, which is given by,

$$\hat{T}^0(\gamma)\Psi(\eta_1,\ldots,\eta_N) = \Psi(\bar{\gamma}\circ\eta_1,\ldots,\eta_n) + \Psi(\eta_1,\bar{\gamma}\circ\eta_2,\ldots,\eta_N) + \Psi(\eta_1,\ldots,\bar{\gamma}\circ\eta_N). \tag{5.35}$$

Notice that this expression only involves wavefunctions of N entries, since due to the Mandelstam identities (3.39) $M_{N+1} = 0$ and consequently $\Psi(\eta_1,\ldots,\eta_{N+1})$ vanishes identically. The fact that we are dealing with a special group (determinant equal 1) implies that

$$\Psi(\eta\circ\eta_1,\eta\circ\eta_2,\ldots,\eta\circ\eta_N) = \Psi(\eta_1,\ldots,\eta_N) \tag{5.36}$$

and therefore by considering $\eta = \bar{\eta}_i$ one immediately concludes that the wavefunctions are really only functions of $N-1$ loops. We will continue using wavefunctions with N entries for convenience.

We complete the small T algebra by representing the $\hat{T}^1$ operator,

$$\hat{T}^a(\gamma_x^x)\Psi(\eta_1,\ldots,\eta_N) = \sum_{k=1}^N \oint_{\eta_k} dy^a\delta^3(x-y) \times[\Psi(\eta_1,\ldots,(\eta_k)_o^y\circ\gamma_x^x\circ(\eta_k)_y^o,\ldots,\eta_N) - \tfrac{1}{N}\hat{T}^0(\gamma)\Psi(\eta_1,\ldots,\eta_N)]. \tag{5.37}$$

To represent the Hamiltonian, we first recall that the magnetic part is given by equation (5.20). Using the formula for $\hat{T}^0(\gamma)$, (5.35), one can immediately realize the action of the magnetic part of the Hamiltonian,

$$\int d^3x\eta_{ab}\mathrm{Tr}(\hat{\tilde{\mathbf{B}}}^a\hat{\tilde{\mathbf{B}}}^b)\Psi(\eta_1,\ldots,\eta_n) = -\tfrac{1}{4}\sum_{i=1}^N\int d^3x\eta^{ac}\eta^{bd}\Delta_{ab}^{(i)}(x)\Delta_{cd}^{(i)}(x)\Psi(\eta_1,\ldots,\eta_n), \tag{5.38}$$

where the loop derivatives $\Delta_{ab}^{(i)}(x)$ are path independent and act only on the ith entry of the loop.

For the electric part, we need to study the action of the operator

$$\hat{\mathcal{E}} = \int d^3x \mathrm{Tr}(\hat{\tilde{\mathbf{E}}}^a \hat{\tilde{\mathbf{E}}}^b)\eta_{ab} \tag{5.39}$$

on the functionals M_N in the connection representation. To perform this calculation it is useful to consider the following identity,

$$\hat{\mathcal{E}}\hat{T}^0(\eta_1) \times \ldots \times \hat{T}^0(\eta_N) = \sum_{i=1}^{N-1} \Big\{ \hat{T}^0(\eta_1) \times \ldots [\hat{\mathcal{E}}, \hat{T}^0(\eta_i)] \ldots \times \hat{T}^0(\eta_N) + \hat{T}^0(\eta_1) \times \ldots \times \hat{T}^0(\eta_{N-1})\hat{\mathcal{E}}\hat{T}^0(\eta_N) \Big\} \tag{5.40}$$

and recalling that the commutator of the electric part of the Hamiltonian with $\hat{T}^0$ is

$$[\int d^3x \mathrm{Tr}(\hat{\tilde{\mathbf{E}}}^a \hat{\tilde{\mathbf{E}}}^b)\eta_{ab}, \hat{T}^0(\gamma)] = -\oint_\gamma dy^a \hat{T}^b(\gamma_y^y)\eta_{ab} + \tfrac{1}{2}\oint_\gamma dy^a \oint_\gamma dy'^b \eta_{ab}\delta^3(y-y')[\hat{T}^0(\gamma_y^{y'})\hat{T}^0(\gamma_{y'}^y) - \tfrac{1}{N}\hat{T}^0(\gamma)] \tag{5.41}$$

and

$$\hat{\mathcal{E}}\hat{T}^0(\gamma) = \tfrac{1}{2}\oint_\gamma dy^a \oint_\gamma dy'^b \eta_{ab}\delta^3(y-y')[\hat{T}^0(\gamma_y^{y'})\hat{T}^0(\gamma_{y'}^y) - \tfrac{1}{N}\hat{T}^0(\gamma)]. \tag{5.42}$$

one can construct explicitly the action of the Hamiltonian constraint in the $SU(N)$ case. Very little is known about this operator in the continuum though some progress has been made in the $SU(3)$ case in the lattice [95] and in the $SU(N)$ case in $1+1$ dimensions [96].

Representations in terms of multiloops also appear in the context of general relativity coupled to gauge fields [97].

5.4 Wilson loops and some ideas about confinement

In his pioneering work, which stimulated most of the interest in the use of loop variables in the treatment of non-Abelian gauge theories, Wilson [48] introduced the idea that the trace of the holonomy could act as an order variable for the theory and could therefore be used to study phase transitions.

The intuitive picture behind this is the following. Consider a Yang–Mills theory coupled to fermions (quarks) and consider the creation and subsequent annihilation of a quark–antiquark pair. Assuming the usual interaction term in the Hamiltonian of the type $\vec{\mathbf{J}} \cdot \vec{\mathbf{A}}$, and neglecting

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

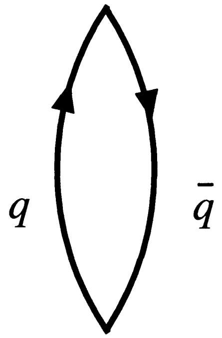

Fig. 5.1. Creation of quark–antiquark pairs viewed as Wilson loops. Creation of a free pair is suppressed if the expectation value of the Wilson loop is a decreasing function of the area of the loop

vacuum polarization effects, one expects such a process to have a weight proportional to the holonomy of the connection $\mathbf{A}_a$ along the closed path formed by the quark–antiquark creation and annihilation process. There are other weight factors independent of the connection and also a weight factor given by the free action of the field.

In order that quarks exist as separate final-state particles it must be possible to consider quark–antiquark processes in which the quark and antiquark lines are well separated, at least when the points of creation and annihilation are far apart. The behavior of the expectation value of the Wilson loop under the separation of the quark–antiquark lines will therefore determine if it is possible for quarks to exist as final states.

For instance, if the expectation value of the Wilson loop turns out to go as $\exp(-l)$, where l is the length of the loop (perturbative and lattice calculations suggest this result for non-confining theories) one sees that one could separate the quark and antiquark lines at will without increasing the length of the loop. This implies that creating well separated quark–antiquark pairs is as likely as creating pairs close together. Therefore the theory is not confining. For instance, for QED an explicit calculation can be performed and $< W >= \exp[-\oint_\gamma dy^\mu \oint dz^\nu D_{\mu\nu}(y-z)]$ where the loop γ is four-dimensional and $D_{\mu\nu}$ is the free propagator of the theory. If one regularizes the calculation one can see that this is proportional to $\exp(-l)$ with l the length of the loop (for details of the regularization see reference [62]).

On the other hand, if $< W >\sim \exp(-a)$, where a is the area of the loop, a process with the quark–antiquark lines shown in figure 5.1 far away from each other is suppressed with respect to one in which the lines are close together and therefore the theory exhibits confinement.

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

These qualitative considerations have been extensively verified in the lattice. It is immediate to confirm them in strong coupling expansions [48, 74, 98], in Monte Carlo simulations [98] and in perturbative calculations.

We therefore see that quark confinement can be thought of as the appearance of confinement of Wilson loops. Since we have extensively argued that Wilson loops can be thought of as lines of electric flux, this gives an image of quark confinement in which lines of electric flux are confined. This is reminiscent of what happens in superconductivity, except that in that case, the confinement refers to lines of magnetic flux. It seems therefore that a system can have two possible confining regimes, one electric and the other magnetic. Each confining regime will be characterized by an order parameter. We argued above that the Wilson loop acted as an order parameter for electric confinement. What could such a parameter be for magnetic confinement? We will now discuss a proposal by t'Hooft for such a parameter and its implications for the loop representation.

t'Hooft [99] introduced a quantity that can be viewed as an order (actually he refers to it as a disorder) parameter for a Yang–Mills theory. The idea can be illustrated by means of the following example in $(2+1)$ dimensions.

Consider an $SU(N)$ Yang–Mills theory in $2+1$ dimensions coupled to an $SU(N)$ Higgs field such that the gauge symmetry is spontaneously and completely broken. Both the Yang–Mills connection $\mathbf{A}_a$ and the Higgs field $\mathbf{H}(x)$ are invariant under gauge transformations generated by the center of the group $SU(N)$, $Z(N)$. A generic element of $Z(N)$ is given by $\exp(2\pi in/N)$ with n an integer. A system like this admits a classical solitonic solution of the following kind. Consider a region R in two-dimensional space surrounded by a region B. In region B symmetry is spontaneously broken, the Higgs field having acquired a "constant" non-zero value. Being an element of the group, "constant" means that there exists a gauge transformation at each point that relates the value of the fields to a certain fixed value, $\mathbf{H}_\Omega(x) = \mathbf{\Omega}(x)\mathbf{H}_0\mathbf{\Omega}(x)^{-1}$. Consider a closed curve in B that surrounds R. Since B is not simply connected, it could happen that by going around the curve, $\mathbf{\Omega}$ becomes multivalued, i.e., $\mathbf{\Omega}_{2\pi} = \exp(i2\pi n/N)\mathbf{\Omega}_0$. We say that the field has a winding number n in such a configuration. The presence of this multivaluedness in the field implies that the configuration is stable. If it were not, it could be radiated away, the final configuration would have $n = 0$ and this could not be achieved from a state with given n in a continuous fashion.

Let us now consider an operator that, starting from a regular field configuration a configuration (such that there exists everywhere a single-valued gauge transformation that maps the field to a constant), will create a configuration like the one we discussed above. To simplify, we will shrink the region R to a single point, at which the gauge transformation mapping the field to a constant is singular. Let us call such a point x_0.

We now define an operator $\phi(x_0)$ that materializes a singular gauge transformation that changes the winding number of the fields,

$$\hat{\phi}(x_0)\Psi[A,H] = \Psi[A_\Omega(x_0), H_\Omega(x_0)], \tag{5.43}$$

where $\Omega(x_0)$ the gauge transformation singular at x_0 such that for every oriented curve $c(\theta)$ that surrounds x_0 once

$$\Omega(x_0)_{\theta=2\pi} = \Omega(x_0)_{\theta=0}\exp(2\pi i/N). \tag{5.44}$$

Let us now consider a state in the physical space of states $\Psi_{ph}(A,H)$. Such states are gauge invariant under regular gauge transformations. Under the singular transformations we are considering here

$$\hat{\phi}(x)\hat{\phi}(y)\Psi_{ph}[A,H] = \hat{\phi}(y)\hat{\phi}(x)\Psi_{ph}[A,H]. \tag{5.45}$$

This statement is self-evident: the resulting gauge transformation only depends on the singularity structure. If one takes a curve surrounding x it will detect the multivaluedness induced by $\hat{\phi}(x)$ and similarly for a curve surrounding y. A curve that surrounds both singularities will detect the combined winding number. All this is independent of the order in which the singularities were added.

The point of this construction was to introduce the operator $\hat{\phi}(x)$. We will now show that this operator plays the role we wanted: that of a "disorder" parameter for the theory. In order to see this, let us study the commutation relation of this operator with the Wilson loop. Acting on a physical state

$$\hat{\phi}(x)W_\gamma[A]\Psi[A,H] = W_\gamma[A_\Omega(x)]\Psi[A_\Omega(x), H_\Omega(x)] \tag{5.46}$$

and noting that $W_\gamma[A_\Omega(x)] = \exp(2\pi i n(\gamma)/N)W_\gamma[A]$, where $n(\gamma)$ is the number of times γ winds around x_0, and $\Psi[A_\Omega(x), H_\Omega(x)] = \phi(x)\Psi[A,H]$ we get

$$\hat{\phi}(x)W_\gamma[A] = \exp(2\pi i n(\gamma)/N)W_\gamma[A]\hat{\phi}(x). \tag{5.47}$$

Let us now consider a basis $|\phi>$ in which the operator $\phi(x)$ is diagonal. In such a basis, the operator $W_\gamma[A]$ introduces a jump of magnitude $\exp(2\pi i n(\gamma)/N)$ in the operator $\phi(x)$ if the point x is within γ. This implies that the Wilson loop acts as a creation operator for a domain inside of which the operator $\phi(x)$ has a different value. Using the natural association of Wilson loops with lines of electric field one can view the domain in which $\phi(x)$ jumps in value as delimited by a closed line of electric field.

This argument can be extended to the case of a theory coupled to fermions (quarks) and in this case one should consider an operator built

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

with a holonomy along an open path with quarks at the ends. By reasoning analogous to that above one can view this open path as a confined line of electric field joining the quark–antiquark pair. Trying to separate the pair requires that the line of confined electric flux be stretched and since the flux is constant, the energy needed to separate the pair is proportional to the distance between the particles. This is a signal of confinement in the theory.

Let us now outline how to generalize the above reasoning to $3+1$ dimensions. In this case the point x_0 at which the gauge transformation was singular becomes a closed line η. Any gauge transformation along a curve γ that is linked with η will be multivalued. The order of multivaluedness is related to the linking number of both curves.

The commutation relation in three dimensions between the Wilson loop and the generalization of $\phi(x_0)$ to three dimensions (which is usually referred to as the t'Hooft operator $B(\eta)$ is

$$W_\gamma B_\eta = B_\eta W_\gamma \exp\left(\frac{2\pi i}{N} GL(\gamma,\eta)\right) \tag{5.48}$$

where $GL(\gamma,\eta)$ is the Gauss linking number of the two curves. The Bs commute among themselves.

The physical results that arise from this picture are that either W or B can exhibit behavior dependent on the area or the length of the loop. According to the possible combinations, four different phases can be identified for the theory. A physical discussion of the four phases in the context of QCD can be found in reference [99], where it is argued that the only relevant phases in the case of pure gauge theories (no fermions) are either electric or magnetic confinement.

The phase in which electric field lines are confined is called the confining phase. From an energetic point of view, this phase is characterized by a degeneracy of the vacuum. This is due to the fact that the Hamiltonian commutes with the operator B and therefore it does not cost extra energy to add magnetic field lines. Electric field lines carry an energy proportional to their length.

The explicit form of the B operator in the connection representation is complicated. t'Hooft [99] was able to find an explicit form for this operator in the lattice and Mandelstam [100] discussed its form in the continuum case. It is remarkable that in the loop representation these operators can be realized in a rather straightforward manner [13].

Recalling the action of the Wilson loop on a state of an $SU(N)$ Yang–Mills theory in the loop representation,

$$\begin{aligned}\hat{T}^0(\gamma)\Psi(\eta_1,\ldots,\eta_N) = \Psi(\bar\gamma\circ\eta_1,\ldots,\eta_n) + \Psi(\eta_1,\bar\gamma\circ\eta_2,\ldots,\eta_N)\\ +\Psi(\eta_1,\ldots,\bar\gamma\circ\eta_N),\end{aligned} \tag{5.49}$$

https://doi.org/10.1017/9781009290203.007 Published online by Cambridge University Press

we can define the operator $\hat{B}(\gamma)$,

$$\hat{B}(\gamma)\Psi(\eta_1,\ldots,\eta_N) = \exp\left(\frac{2\pi i}{N}\sum_{k=1}^{N} GL(\gamma,\eta_k)\right)\Psi(\eta_1,\ldots,\eta_N), \quad (5.50)$$

where

$$GL(\gamma,\eta_k) = \frac{1}{4\pi}\oint_\gamma dx^a \oint_{\eta_k} dy^b \epsilon_{abc}\frac{(x-y)^c}{|x-y|^3} \quad (5.51)$$

is the Gauss linking number of γ and η_k. This topological invariant measures how many times the loop η_k "threads through" the loop γ (for more details see chapter 10).

It is straightforward to study the commutation relations of this operator with the Yang–Mills Hamiltonian. Because the Gauss linking number is a topological invariant, it commutes with the portion of the Hamiltonian with two loop derivatives, since adding an infinitesimal loop does not change the value of the topological invariant. This point really requires a regularization since the Hamiltonian adds an infinitesimal loop at all points in the manifold and could introduce divergences. The electric part of the Hamiltonian also commutes with the $\hat{B}$ operator. This can be seen by recalling that the effect of the electric part of the Hamiltonian on a wavefunction of N loops is to produce a wavefunction with $N+1$ loops produced by fissions of loops at their self-intersections, as shown in equations (5.28),(5.42). Computing the linking number before or after the fission gives the same result and the operators commute. The reader may be interested in what happens if one characterizes the action of the Hamiltonian purely in terms of N loops using the Mandelstam identity, as we did in the $SU(2)$ case, equation (5.28). In this case some portions of the loop are rerouted and some of the linking numbers — which depend on the orientation — may change sign. However, the result remains unchanged, because the operator $\hat{B}$ takes values on $Z(N)$ and this makes the operator compatible with the Mandelstam identities.

So we see that the operator $\hat{B}$ commutes with the Hamiltonian. We will now study the commutation relation of the Wilson loop with the Hamiltonian. In the loop representation the Wilson loop becomes the $\hat{T}^0(\gamma)$ operator and let us assume that we are considering loops γ that are smooth. Consider the commutator of the $\hat{T}^0(\gamma)$ operator with the electric part of the Hamiltonian, equation (5.42). Its action on a state in the loop representation can be computed using (5.35),(5.37). From (5.42) one can see that there are two terms. The first one is proportional to the length of the loop (double integral along γ) and the second one, taking into account (5.37), involves an integral along γ and another integral along the loop that appears in the argument of the wavefunction. If one considers long loops, the first term (proportional to the length) dominates the other term, which involves intersections of γ with the argument of the wavefunction. The commutator of $\hat{T}^0(\gamma)$ with the magnetic part of the Hamiltonian vanishes, which can immediately be seen from (5.38) since the loop derivatives act specifically on the arguments of the wavefunction and the loop dependence of the $\hat{T}^0$ is transparent.

We therefore see that it is simple to prove that the loop representation naturally describes the confining phase of Yang–Mills theory. There is a natural representation of the disorder operator $\hat{B}$ and adding an electric field line has an energetic cost proportional to the length of the line being added, which is one of the signs of confinement. It should be remarked that these arguments show that there exists a confining phase in which it is energetically expensive to create Wilson loops. However, one could conceive different phases, where the distribution of loops is dense in space and then the dominating term in the expressions considered above, instead of being that of the length of the loop could be the one involving intersections. However, the fact that $\hat{B}$ *always* commutes with the Hamiltonian suggests that Yang–Mills is always in a confining phase.

5.5 Conclusions

We have considered various loop-based approaches to Yang–Mills theories. We have emphasized the use of Hamiltonian techniques and the loop representation, which we constructed explicitly for $SU(2)$ and $SU(N)$ Yang–Mills theories. As the reader may have perceived, the treatment of Yang–Mills theories in the language of loops in the continuum has only a formal character and little progress has actually been made towards understanding the non-perturbative physics of QCD. Only qualitative arguments, like the ones we introduced in the previous section, shed some light on the various physical processes of non-Abelian gauge theories. On the other hand, the gauge invariant description of Yang–Mills theories based on holonomies has found application in attempts to set the theory in a more mathematically rigorous basis. For instance, it may be possible to define an infinite-dimensional measure rigorously in the space of connections modulo gauge transformations in terms of the loop algebra [40]. Progress in this respect has also been made in lower dimensions [96, 102, 103]. In this chapter we have concentrated on pure Yang–Mills theories. Coupling to fermions and Higgs fields can be introduced in the loop representation but again most results are only formal. In the next chapter we will return to the issue of matter couplings in Yang–Mills theories.