SU(*N*) Quantum Yang−Mills theory in two dimensions: A complete solution

Abhay Ashtekar
Center for Gravitational Physics and Geometry, Physics Department, Penn State, University Park, Pennsylvania 16802-6300

Jerzy Lewandowski
Instytut Fizyki Teoretycznej, Uniwersytet Warszawski, ul. Hoża 69, 00-681 Warsaw, Poland

Donald Marolf
Department of Physics, Syracuse University, Syracuse, New York 13244-1130

José Mourão
Departamento de Fisica, Instituto Superior Tecnico, Av. Rovisco Pais 1, 1096 Lisboa Codex, Portugal

Thomas Thiemann
Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Schlaatzweg 1, 14473 Potsdam, Germany

A closed expression of the Euclidean Wilson-loop functionals is derived for pure Yang−Mills continuum theories with gauge groups SU(*N*) and U(1) and space-time topologies $\mathbb{R}^1 \times \mathbb{R}^1$ and $\mathbb{R}^1 \times S^1$. (For the U(1) theory, we also consider the $S^1 \times S^1$ topology.) The treatment is rigorous, manifestly gauge invariant, manifestly invariant under area preserving diffeomorphisms and handles all (piecewise analytic) loops in one stroke. Equivalence between the resulting Euclidean theory and and the Hamiltonian framework is then established. Finally, an extension of the Osterwalder−Schrader axioms for gauge theories is proposed. These axioms are satisfied in the present model. © *1997 American Institute of Physics.* [S0022-2488(97)00911-0]

I. INTRODUCTION

Although the literature on Yang−Mills theories in 2 space−time dimensions is quite rich, a number of issues have still remained unresolved. The purpose of this paper is to analyze three such issues. The paper is addressed both to high energy theorists and mathematical physicists. Therefore, an attempt is made to bridge the two sets of terminologies, techniques and conceptual frameworks.

The first issue concerns the expectation values of traces of holonomies of the connection around closed loops in the Euclidean domain, i.e., the Wilson loop functionals. The traces of holonomies are, arguably, *the* central observables of the (pure) Yang−Mills theory. In the classical regime, they constitute a natural set of (over)complete gauge invariant functions of connections with rich geometrical and physical content. Hence, their Euclidean vacuum expectation values are the natural *gauge invariant* analogs of the expectation values $\chi(f) := \langle \exp i \int d^n x \, \phi(x) f(x) \rangle$ in scalar field theories which determine all the n-point (i.e., Schwinger) functions (via repeated functional differentiation with respect to f). From theoretical physics considerations, therefore, one expects the Wilson loop functionals to completely determine the theory. From a mathematical physics perspective, the quantum theory is completely determined if one specifies the underlying measure $d\mu$—the rigorous analog of the heuristic expression ''exp$-S[A]\mathscr{D}A$'' — on the space of Euclidean paths. The expectation values of products of traces of holonomies determine the ''moments'' of the measure $d\mu$. Hence, one expects them to determine the measure completely.

Over the years, these considerations inspired a number of authors to devise imaginative ways to explore properties of the Wilson loop functionals. For example, Makeenko and Migdal[1] formulated differential equations that these functions have to satisfy *on the space of loops* and then introduced physically motivated ansätze to solve them. (See also Refs. 2, 3.) Similarly, Gross and co-authors[4] have used stochastic methods to obtain closed expressions for non-overlapping Wilson loops. While these methods have yielded a wealth of insights, to the best of our knowledge, a closed expression for generic Wilson loops has not yet appeared in the literature. (At best the compuations performed provide us with an expression which is exact but only implicit in the sense that there have still to be done non-trivial computations for each case at hand; see, e.g., Ref. 5). The first purpose of this paper is to provide such an expression for SU(*N*) [and U(1)] gauge theories assuming that the underlying Euclidean space–time has a topology of $\mathbb{R}^1 \times \mathbb{R}^1$, or $\mathbb{R}^1 \times S^1$. (In the U(1) case, we also allow the topology to be $S^1 \times S^1$.) The final expression is explicit up to a trivial contraction of group indices for a matrix which we have computed for the general case.

The second issue treated here is the relation between the Euclidean description in terms of functional integrals and the canonical description in terms of a Hilbert space and a Hamiltonian. For scalar field theories, there exists a general framework that ensures this equivalence (see, e.g., Ref. 6). We extend it to gauge theories and explicitly establish the equivalence between the two descriptions in the case when the Euclidean topology is $\mathbb{R}^1 \times \mathbb{R}^1$ or $\mathbb{R}^1 \times S^1$. While the extension involved is rather straightforward, it is quite illuminating to see how the Euclidean framework— which, *a priori*, does not know that the system has only a finite number of true degrees of freedom—reduces to the Hamiltonian framework which, from the very beginning, exploits the fact that this is a quantum mechanical system, disguised as a quantum field theory.

Our third goal is to suggest an extension of the axiomatic framework of Osterwalder and Schrader. In that framework, one assumes from the very beginning that the underlying space of paths is linear, and can be identified with the distributional dual $\mathscr{S}'$ of the Schwartz space $\mathscr{S}$ of smooth test functions of rapid decrease (see, e.g., Ref. 6). The axioms are restrictions on the measure μ on $\mathscr{S}'$, formulated as conditions on the functional $\chi(f) = \int d\mu (\exp i \int d^n x\, \phi(x) f(x))$, introduced above, now interpreted as the Fourier transform of the measure μ. Now, in gauge theories, it is natural to regard each gauge equivalence class of connections as a distinct physical path. The space $\mathscr{A}/\mathscr{G}$ of paths is then a genuinely non-linear space and the standard axioms can not even be stated unless one introduces, via gauge fixing, an artificial linear structure on $\mathscr{A}/\mathscr{G}$. (In higher dimensions, due to Gribov ambiguities, such a gauge fixing does not exist.) We will suggest a possible extension of the standard framework to encompass gauge theories in a manifestly gauge invariant fashion and show that the axioms are in fact satisfied in the two-dimensional Yang–Mills theories discussed in sections II–IV. We would like to emphasize, however, that there is a key difference between the status of the first two sets of results and the third. In the first two cases, we deal only with two-dimensional Yang–Mills theory and the results are definitive. In the third part, the general framework is applicable to gauge theories in any space–time dimension and the discussion is open-ended; it opens a door rather than closing one. In particular, relative to other attempts[7] in the literature, our approach is still very much in the preliminary stage.

The main ideas behind our approach can be summarized as follows. (For a more detailed discussion, see Refs. 8, 9.) First, we will maintain manifest gauge invariance in the sense that we will work directly on the space $\mathscr{A}/\mathscr{G}$. No attempt will be made to impose a vector space structure by gauge-fixing; we will face the non-linearities of $\mathscr{A}/\mathscr{G}$ squarely. Now, it is well-known that, in quantum field theory, smooth fields make a negligible contribution to the path integrals; physically interesting measures tend to be concentrated on distributions. Therefore, in the case of gauge theories, we need to allow generalized connections. Fortunately, a suitable completion $\overline{\mathscr{A}/\mathscr{G}}$ of the space $\mathscr{A}/\mathscr{G}$ of smooth physical paths has been available in the literature for some time.[10,11] Furthermore, this space carries[11–13] a rigorously defined, uniform measure μ_0 which can serve as a fiducial measure—the analog of the heuristic measure $\mathscr{D}A$. The idea is to construct the physically relevant measure by "multiplying $d\mu_0$ by $\exp{-S}$," where S is the Yang–Mills action.

As in all constructive quantum field theories, this task is, of course, highly non-trivial. We proceed in the following steps. First, we consider Wilson's lattice-regularized version S_W of S. Now, it turns out that $\exp{-S_W}$ *is* an integrable function with respect to the measure $d\mu_0$ and, furthermore, products of traces of holonomies, $T_{\alpha_1} \cdots T_{\alpha_k}$, around loops $\alpha_1, \ldots, \alpha_k$ are integrable on $\overline{\mathscr{A}/\mathscr{G}}$ with respect to the measure $\exp{-S_W} d\mu_0$. We compute these expectation values as a function of the lattice spacing, used in the Wilson regularization, and then show that the resulting expressions have a well-defined limit as the spacing goes to zero. These are the required Wilson-loop functionals in the continuum. General theorems[10–12] from integration theory on $\overline{\mathscr{A}/\mathscr{G}}$ guarantee that there exists a genuine, normalized measure μ_{YM} on $\overline{\mathscr{A}/\mathscr{G}}$ such that the integrals of products of traces of holonomies with respect to μ_{YM} are the Wilson loop functionals computed by the regularization procedure. This provides a concrete proof of the existence of a consistent Euclidean theory.

The techniques we use were first developed in the context of a non-perturbative approach to general relativity.[14] Therefore, our emphasis is often different from that in the literature of Yang–Mills theories. For instance, we arrive at the final, closed expressions of Wilson loops by a direct computation of the functional integrals, rather than through differential equations these functionals satisfy on the loop space. In this sense, our approach is similar to that followed in the mathematical physics literature. However, in these rigorous approaches, one often tries to exploit methods which have been successful in kinematically linear theories and, to do so, introduces a vector space structure of $\mathscr{A}/\mathscr{G}$ through gauge fixing. As mentioned above, we work directly on the non-linear space $\mathscr{A}/\mathscr{G}$ and thus avoid gauge fixing in conceptual considerations. Also, our method respects the invariance of the theory under area preserving diffeomorphisms. In particular, our Wilson loop functionals — and hence the final, physical measure for the continuum theory — are manifestly invariant under the action of this group.

The plan of the paper is as follows. In section II, we review the relevant notions from calculus on $\mathscr{A}/\mathscr{G}$. In section III, we reformulate lattice gauge theory in a manner that makes the analytic computation of Wilson loop functionals easier. This formulation constitutes the basis of our discussion of the continuum theory in section IV. Here, we first derive the general form of the Wilson loop functionals with ultraviolet and infrared cut-offs provided by the lattice regularization and then show that the functionals admit well-defined limits as the cut-offs are removed. In the mathematical physics terminology, these limits are the generating functions for the physical, Yang–Mills measure on $\overline{\mathscr{A}/\mathscr{G}}$. For simple loops, we recover the well-known area law which is generally taken to be the signature of confinement. More generally, if we suitably restrict our choices of loops, our general results reduce to those obtained previously in the mathematical physics literature. Section IV reviews the Hamiltonian quantization of Yang–Mills fields in cases when the underlying Lorentzian space–time has the topology of a 2-plane or a 2-cylinder. The aim of section V is threefold. We begin with a brief review of the Osterwalder–Schrader framework for kinematically linear theories and, using the machinery developed in sections II–IV, propose an extension to handle gauge theories. We then show that our two-dimensional model, treated in section IV, satisfies these axioms. Finally, we show that the Hamiltonian framework reviewed in section IV can be systematically recovered from the Euclidean framework. Section VI summarizes the main results, compares them with the results available in the literature and suggests some directions for further work.

A number of technical topics are covered in appendices. Specifically, Young tableaux which are needed in certain computations of Sec. IV are discussed in Appendix A and the details of the Euclidean U(1) theory on a torus are presented in Appendix B.

Finally, we wish to emphasize that in most of this paper we have restricted ourselves to non-compact space–times since it is only in this case that a direct comparison with the Hamiltonian theory is possible. In particular, all our results in the non-Abelian case pertain to space–time topologies $\mathbb{R} \times \mathbb{R}$ and $S^1 \times \mathbb{R}$ and it is only in these cases that we have obtained a complete solution. In recent years, there has been extensive work on Euclidean Yang–Mills theories on

compact Riemannian surfaces (see, e.g., Refs. 15–17) which has led to a variety of interesting results. (For reviews, see, e.g., Refs. 18, 19.) Similarly, interest in the large *N* limit of two-dimensional Yang–Mills theories and coupling to fermions in this limit has also been renewed in the last few years. (See, e.g., Refs. 20–22.) However, all these developments lie outside the scope of the present paper.

II. PRELIMINARIES

In this section, we will review the basic notions from Refs. 10–13, 23 (and references therein) which will be used in this paper. This material will provide the necessary background for our discussion of the mathematical aspects of functional integration, axiomatic formulation of gauge theories and the relation between Euclidean and Hamiltonian formulations. A reader who is interested primarily in the computation of the Wilson loop functionals can skip this material and go directly to sections III and IV.

By a loop we will mean a piecewise-analytic embedding of S^1 into the (Euclidean) space–time manifold M. For technical convenience, we will only consider based loops, i.e., loops passing through a fixed point p in M. Denote the set of these loops by $\mathscr{L}_p$. As indicated already, our structure group will be either SU(N) (where $N \geq 2$) or U(1). Fix any one of these groups, consider a trivial Principal fibre bundle B on M and denote by $\mathscr{A}$ the space of smooth connections on B. Given any $A \in \mathscr{A}$, we can associate with every $\alpha \in \mathscr{L}_p$ an element of SU(N) by evaluating the holonomy,

$$h_\alpha(A) := \mathscr{P} \exp\left(- \oint_\alpha A \right), \tag{II.1}$$

at the base point p (where, as usual, $\mathscr{P}$ stands for ''path ordered''). Let us introduce an equivalence relation on $\mathscr{L}_p$: two loops $\alpha_1, \alpha_2 \in \mathscr{L}_p$ will be said to be holonomically equivalent, $\alpha_1 \sim \alpha_2$, iff $h_{\alpha_1}(A) = h_{\alpha_2}(A) \forall A \in \mathscr{A}$. Each of these holonomically equivalent loops will be called a *hoop*. It is straightforward to verify that the space $\mathscr{HG}$ of hoops has a natural group structure. We will call it the *Hoop group*. For notational simplicity, in what follows we will not distinguish between a hoop and a loop in the corresponding equivalence class.

Denote by $\mathscr{G}$ the group of smooth, local gauge transformations (i.e., of smooth vertical automorphisms of B). Of special interest are the $\mathscr{G}$-invariant functions T_α of connections obtained by taking traces of holonomies:

$$T_\alpha(A) := \frac{1}{N} \operatorname{tr}(h_\alpha(A)), \tag{II.2}$$

where the trace is taken with respect to the *N*-dimensional fundamental representation of the structure group. As is well known, the functions T_α suffice to separate points of $\mathscr{A}/\mathscr{G}$ in the sense that given all the T_α, we can reconstruct the smooth connection modulo gauge transformations.[24] This is significant because, in the classical theory, physical paths are represented by elements of $\mathscr{A}/\mathscr{G}$.

To go over to the quantum theory, we need to extend this space of paths appropriately since the set of smooth paths is, typically, of zero measure in physically interesting theories. One possible extension has been carried out in the literature.[10,11] (For motivational remarks, see Ref. 8.) This extension, $\overline{\mathscr{A}/\mathscr{G}}$, can be characterized in three complementary ways, each emphasizing a different set of its properties. Since $\overline{\mathscr{A}/\mathscr{G}}$ will play a fundamental role in the quantum theory — in our approach it represents the space of gauge invariant, physical paths in the Euclidean approach — we will now sketch all these characterizations:

(i)   Perhaps the simplest characterization is the following: $\overline{\mathscr{A}/\mathscr{G}}$ is the space of *all* homomorphisms from the hoop group $\mathscr{HG}$ to the structure group SU(N) or U(1), (modulo the adjoint action of the structure group at the base point p). It is obvious that, given a smooth connection, the holonomy map of (II.1) provides such a homomorphism. However, it is easy to construct[11] examples of more general homomorphisms which, for example, would correspond to "distributional connections." In relation to the more familiar scalar field theories, $\mathscr{HG}$ will play a role which in some ways is similar to that played by the space $\mathscr{S}$ of test functions and $\overline{\mathscr{A}/\mathscr{G}}$ is analogous to the space $\mathscr{S}'$ of Schwartz distribution. In particular, just as $\mathscr{S}'$ is the space of paths for scalar fields, $\overline{\mathscr{A}/\mathscr{G}}$ will serve as the space of paths for gauge theories. The "duality" between $\mathscr{HG}$ and $\overline{\mathscr{A}/\mathscr{G}}$ is non-linear. However, just as elements of $\mathscr{S}$ serve as labels for cylindrical functions on $\mathscr{S}'$, elements of $\mathscr{HG}$ will serve as labels for cylindrical functions on $\overline{\mathscr{A}/\mathscr{G}}$.

(ii)   The second characterization brings out the topological structure of $\overline{\mathscr{A}/\mathscr{G}}$. Recall first that in any of the standard Sobolev topologies on $\mathscr{A}/\mathscr{G}$, the functions T_α are continuous. Furthermore, for gauge groups under consideration, they are bounded. Hence, the $\star$-algebra they generate is a sub-algebra of the $C^\star$-algebra $C^0(\mathscr{A}/\mathscr{G})$ of all continuous bounded functions on $\mathscr{A}/\mathscr{G}$. Denote the completion of this $\star$-algebra by $\mathscr{HA}$. This is an Abelian $C^\star$-algebra with identity and is called the *holonomy algebra*. Now, the Gel'fand representation theory guarantees that $\mathscr{HA}$ is naturally isomorphic with the $C^\star$-algebra of all continuous functions on a compact Hausdorff space. This space — the Gel'fand spectrum of $\mathscr{HA}$ — is our $\overline{\mathscr{A}/\mathscr{G}}$. Thus, the topology on $\overline{\mathscr{A}/\mathscr{G}}$ is the coarsest one which makes the Gel'fand transforms of the traces of holonomies continuous. Finally, since $\mathscr{HA}$ suffices to separate points of $\mathscr{A}/\mathscr{G}$, it immediately follows that $\mathscr{A}/\mathscr{G}$ is densely embedded in $\overline{\mathscr{A}/\mathscr{G}}$.

(iii)   The last characterization is in terms of projective limits.[25] One begins with two projective families labelled by graphs, each consisting of compact Hausdorff manifolds. The projective limit of the first yields a completion $\overline{\mathscr{A}}$ of the space $\mathscr{A}$ of smooth connections while the projective limit of the second provides a completion of the the space $\mathscr{G}$ of smooth gauge transformations. One then shows that $\overline{\mathscr{A}/\mathscr{G}} = \overline{\mathscr{A}}/\overline{\mathscr{G}}$. This characterization is best suited for analyzing the (surprisingly rich) geometric structure of $\overline{\mathscr{A}/\mathscr{G}}$.[12,26]

Finally, we note that $\overline{\mathscr{A}/\mathscr{G}}$ admits[11–13] a natural, normalized, Borel measure μ_0 which, in our approach, will play the role that "$\mathscr{D}A$" plays in heuristic considerations. We will conclude by indicating how this measure is defined.

To begin with, let us consider the family of all piecewise analytic, oriented graphs Γ in M. Denote by $\pi_1(\Gamma)$ the fundamental group of the graph Γ. Choose a system of generators $\beta_1, \ldots, \beta_n$ of $\pi_1(\Gamma)$ where $n := \dim(\pi_1(\Gamma))$ is the number of independent generators of the fundamental group. With this machinery at hand, we can define the notion of "cylindrical functions," which will be the simplest functions on $\overline{\mathscr{A}/\mathscr{G}}$ that we will be able to integrate. Note first that, given any graph Γ, we have a natural projection map,

$$p_\Gamma : \overline{\mathscr{A}/\mathscr{G}} \to G^n \quad A \to (h_{\beta_1}(A), \ldots, h_{\beta_n}(A)), \tag{II.3}$$

from $\overline{\mathscr{A}/\mathscr{G}}$ to G^n, where G is the structure group [i.e., SU(N) or U(1)] under consideration. Cylindrical functions are obtained by pull-backs of smooth functions on G^n under this map. Thus, given any smooth function f_Γ on G^n, $f = (p_\Gamma)^* f_\Gamma$ is a cylindrical function.

The measure μ_0 on $\overline{\mathscr{A}/\mathscr{G}}$ can now be introduced via:

$$\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu_0(A) f(A) := \int_{G^n} d\mu_H(g_1) .. d\mu_H(g_n) f_\Gamma(g_1, \ldots, g_n). \tag{II.4}$$

The proof that this condition does indeed define an infinite dimensional, (σ-additive) regular, normalized Borel measure μ_0 on $\overline{\mathscr{A}/\mathscr{G}}$ is given in Ref. 23.

III. LATTICE GAUGE THEORY

In this section, we will recast the standard description of lattice gauge theory in a form that is better suited for our discussion of the continuum limit in section IV.

Consider finite square lattices $\Gamma(a,L_x,L_y)$ in M with spacing a and length L_x and L_y in the x and y directions. This lattice contains $(N_x+1)(N_y+1)$ vertices, where $N_x a := L_x$, $N_y a := L_y$. Note that the use of such a lattice for quantum field theory implies both an infra-red regulator (the finite volume defined by the L_x and L_y) and an ultra-violet regulator (defined by the lattice spacing a). Our strategy will be to construct a regulated quantum theory in this section and then remove the regulators in the next section.

Let us denote the open path along an edge (link) of the lattice from a vertex i to an adjacent vertex j by

$$l = l_{i \to j}$$

so that we may define the plaquette loops

$$\Box_{(x,y)} := l^{-1}_{(x,y)\to(x,y+1)} \circ l^{-1}_{(x,y+1)\to(x+1,y+1)} \circ l_{(x+1,y)\to(x+1,y+1)} \circ l_{(x,y)\to(x+1,y)} . \qquad \text{(III.1)}$$

That is, each plaquette loop starts at the bottom left corner and our convention is such that the coordinate directions define positive orientation. Here the coordinates x,y are taken to be integers. For the plane $M = \mathbb{R} \times \mathbb{R}$, all of these links are distinct while for the cylinder, $M = \mathbb{R} \times S^1$, we identify $l_{(1,y)\to(1,y+1)} \equiv l_{(N_x+1,y)\to(N_x+1,y+1)}$. On the torus, we also identify $l_{(x,1)\to(x+1,1)} \equiv l_{(x,N_y+1)\to(x+1,N_y+1)}$.

Next, we introduce a set of closed loops which can serve as generators, i.e., in terms of which any loop in Γ based at p can be expressed via composition:

(i) Let $\rho_{x,y}$ be an open path in Γ from p to the point (x,y). The loops

$$\beta_{x,y} := \beta_{\Box_{(x,y)}} := \rho^{-1}_{x,y} \circ \Box_{(x,y)} \circ \rho_{x,y} \qquad \text{(III.2)}$$

generate all loops on the plane.

(ii) On the cylinder, we need an additional loop. We will take it to be the ''horizontal'' loop

$$\gamma_x := l_{(N_x,N_y+1)\to(1,N_y+1)} \circ l_{(N_x-1,N_y+1)\to(N_x,N_y+1)} \circ \cdots \circ l_{(1,N_y+1)\to(2,N_y+1)} . \qquad \text{(III.3)}$$

(iii) Similarly, on the torus we need an additional loop,

$$\gamma_y := l_{(1,N_y)\to(1,1)} \circ l_{(1,N_y-1)\to(1,N_y)} \circ \cdots \circ l_{(1,1)\to(1,2)} . \qquad \text{(III.4)}$$

However, the loops $\{\beta_{x,y}, \gamma_x, \gamma_y\}$ are not independent as the loop $\gamma_y^{-1} \circ \gamma_x^{-1} \circ \gamma_y \circ \gamma_x$ can be written as a composition of the $\beta_{x,y}$. (An intuitive notion of independence will suffice for our work here. For a careful definition, see Ref. 11.) This constraint will lead to an ''interacting'' U(1) theory for the torus in contrast to the plane and the cylinder.

With these preliminaries out of the way, let us now summarize the standard formulation of the lattice gauge field theory by Wilson.[27] For each of the links in the lattice, introduce one G-valued degree of freedom (the ''parallel transport along the link''). Let the ''lattice Yang–Mills action'' be given by the Wilson expression

$$S_W := \sum_{\Box} \left[1 - \frac{1}{N} \Re \operatorname{tr}(h_\Box) \right], \qquad \text{(III.5)}$$

where $h_\square$ denotes product of link variables around the plaquette $\square$ and $\mathfrak{R}$ tr is the real part of the trace. Also, let $d\mu_W$ be the Haar measure on G^{N_l}, where N_l is the number of links in the graph. The regulated Wilson-loop functional is now given by

$$\langle T_{\alpha_1} \cdots T_{\alpha_k} \rangle := \frac{1}{Z(a;L_x,L_y)} \int_{G^{N_l}} d\mu_W \, e^{-\beta S_W} T_{\alpha_1} \cdots T_{\alpha_k}, \tag{III.6}$$

where $\alpha_1, \ldots, \alpha_k$ are loops in $\Gamma = \Gamma(a;L_x,L_y)$; the ''inverse temperature'' is given by

$$\beta = \frac{1}{g_0^2 a^{4-d}} \tag{III.7}$$

($d=2$ being the dimension of M); and where $g_0 = g_0(a)$ is the bare coupling constant. The partition function $Z = Z(a;L_x,L_y)$ is defined through $\langle T_p \cdots T_p \rangle = 1$ where p denotes the trivial loop at p. From a mathematical physics perspective, these Wilson loop functionals can also be regarded as the characteristic functional of the regulated measure. To emphasize this dual interpretation, using the standard notation for characteristic functionals, we will set:

$$\chi(\alpha_1, \ldots, \alpha_k; a;L_x,L_y) := \langle T_{\alpha_1} \cdots T_{\alpha_k} \rangle. \tag{III.8}$$

For our purposes, it will turn out to be more convenient to re-express the characteristic functional in terms of integrals over the independent *loops* in the graph Γ. To do so, we make use of the fact that, whenever it is used to integrate gauge invariant functions, the measure $d\mu_W$ may be replaced by the Haar measure on G^N, where N is the number of independent loop generators of the graph Γ. This fact follows immediately from the results of Refs. 13, 26. (In the language of these works, it is contained in the statement that $\overline{\mathscr{A}/\mathscr{G}} = \overline{\mathscr{A}/\mathscr{G}}$ and that the Haar measure on $\overline{\mathscr{A}}$ projects unambiguously to yield the Haar measure on $\overline{\mathscr{A}/\mathscr{G}}$.) Thus, we may write the regulated characteristic functional as:

$$\chi(\alpha_1, \ldots, \alpha_k) = \frac{1}{Z} \int_{G^N} \prod_\square d\mu_H(g_\square) \exp(-\beta S_W) \begin{cases} \displaystyle\prod_{i=1}^{k} \, \text{tr} \, \alpha_i(g_\square) & : \text{on} \ \mathbb{R}^2 \\[2ex] \displaystyle\int_G d\mu_H(g_x) \prod_{i=1}^{k} \, \text{tr} \, \alpha_i(g_\square,g_x) & : \text{on} \ \mathbb{R}^1 \times S^1 \end{cases}, \tag{III.9}$$

where $d\mu_H$ is the Haar measure on G and $\alpha_i(g_\square)$ is the expression for α_i in terms of the generators $\beta_{x,y}$ with each generator $\beta_{x,y}$ replaced by the integration variable $g_{x,y}$ [similarly for $\alpha_i(g_\square,g_x)$]. The corresponding expression for the torus will appear at the end of this section. The idea of the next section will simply be to evaluate the above integrals for any given a,L_x,L_y and then take the limits to remove the ultra-violet and infra-red regulators.

To conclude this section, we will introduce some definitions and collect a few facts about loops in Γ. These will be useful in section IV.

Definition III.1: A loop is said to be simple iff there is a holonomically equivalent loop which has no self-intersections.

Note that any simple loop which is homotopically trivial (in space–time) divides space–time into two regions: an interior which is topologically a 2-disk and an exterior. This is just the Jordan curve theorem.

Definition III.2: On the torus, we define the surface enclosed by a simple homotopically trivial loop to lie on the left as one follows the loop counterclockwise (when the torus is represented as a two-dimensional rectangle with the standard identifications.)

Definition III.3: Two distinct simple homotopically trivial loops are said to be non-overlapping iff the intersection of the surfaces that they enclose has zero area. The homotopically non-trivial loops β_x and β_y will both be said not to overlap any other loop.

So, for example, all the loops $\beta_{x,y}$ are simple since they lie in the same hoop class as the plaquette loops $\square_{(x,y)}$. Non-overlapping distinct simple loops are allowed to share whole segments whence the plaquette generators of our graph (lattice) are mutually non-overlapping.

It will turn out that the following two simple lemmas govern the form of the characteristic functional in two space–time dimensions.

Lemma III.1: Every simple, homotopically trivial loop α on Γ can be written as a particular composition of the generators $\beta_\square$ contained in the surface enclosed by α, with each $\beta_\square$ appearing once and only once.

It is readily checked that when two homotopically trivial loops α_1 and α_2 (enclosing disks D_1 and D_2) are non-overlapping and such that $D_1 \cup D_2$ is also a disk, then either $\alpha_1\alpha_2$ or $\alpha_1\alpha_2^{-1}$ (or, on the torus, perhaps the inverse of one of these loops) encloses $D_1 \cup D_2$. Since every disk is a finite union of plaquettes, Lemma III.1 follows immediately. $\square$

This Lemma allows us to write a simple expression for the generating functional on the torus. Note that, after 'ungluing' the torus to make a rectangle, the loop $\gamma_x \circ \gamma_y \circ \gamma_x^{-1} \circ \gamma_y^{-1}$ is simple and homotopically trivial, enclosing the entire area of the torus. As a result, it may be written as a product of the plaquette loops $\beta_\square$ in which each $\beta_\square$ appears once and only once. We may therefore pick any one of these loops (say $\beta_{(0,0)}$) and write it as a function of the other plaquette loops and the loops γ_x, γ_y. Alternatively, we find a product C of holonomies along all the loops $\beta_\square, \gamma_x, \gamma_y$ which is the identity in G. Inserting a δ distribution on G enforcing the constraint $C = 1_N$ we find for the generating functional on the torus

$$\chi(\alpha_1, \ldots, \alpha_k) = \frac{1}{Z} \int_{G^N} \prod_\square d\mu_H(g_\square) d\mu_H(g_x) d\mu_H(g_y) \exp(-\beta S_W) \prod_{i=1}^{k} \text{tr } \alpha_i(g_\square, g_x, g_y)$$

$$= \frac{1}{Z} \int_G \prod_\square d\mu_H(g_\square) d\mu_H(g_x) d\mu_H(g_y) \delta(C, 1_N) \exp(-\beta S_W)$$

$$\times \prod_{i=1}^{k} \text{tr } \alpha_i(g_\square, g_x, g_y). \tag{III.10}$$

Finally, we have:

Lemma III.2: Every loop can be written as a composition of simple non-overlapping loops.

This follows from the fact that the $\beta_\square$ (together with β_x, β_y on $S^1 \times \mathbb{R}$ and T^2) are simple and non-overlapping and that they generate the graph Γ. $\square$

IV. CONTINUUM THEORY

In this section we will derive a closed expression for the Wilson loop functionals — i.e., for the characteristic functional of the measure — for the continuum theory when the underlying manifold M is either a 2-plane or a cylinder. (For the torus, we have been able to carry out the computation to completion only for the Abelian case, $G = U(1)$, and this theory is discussed in detail in Appendix B.)

In section IV A, we will discuss U(1) theories and in section IV B, SU(*N*) theories. In both cases, we will show that the lattice-regulated characteristic functional admits a well-defined limit as the ultra-violet and infrared cut-offs are removed. Furthermore, we will be able to read-off certain qualitative properties of these functionals. However, the explicit expression involves a group-dependent constant. This is evaluated in section IV C.

A. Abelian case (U(1))

Let us first note that, in the U(1) case, products of functions T_α can be reduced to a single $T_{\alpha'}$ in the obvious fashion. Therefore, we need to consider only single loops. Fix a loop α and consider its decomposition into non-overlapping simple loops. Let k_I be the effective winding number of the simple homotopically trivial loop α_I, $I = 1, \ldots, n$ and let l_x, l_y be winding numbers of the homotopically non-trivial loops β_x, β_y in this decomposition. Define $|\alpha_I|$ to be the number of plaquettes enclosed by the simple loop α_I. We can then write the characteristic functional as follows [with $G = $U(1)]:

$$\chi(\alpha) = \frac{1}{Z} \int \prod_\square \left[\int_G d\mu_H(g_\square) \exp(-\beta(1 - \Re(g_\square))) \prod_{I=1}^n \left(\prod_{\square \in \alpha_I} g_\square \right)^{k_I} \right]$$
$$\times \int_G d\mu_H(g_x) g_x^{l_x} \int_G d\mu_H(g_y) g_y^{l_y} \delta\left(\prod_\square g_\square, 1 \right)^e,$$

where we could neglect the precise ordering of plaquette variables (that occurred in the decomposition of α in terms of $\beta_\square, \beta_x, \beta_y$) because the gauge group is Abelian. In this formula $l_x = l_y = 0$ on the plane and $l_y = 0$ on the cylinder and $e = 0$ for the plane and the cylinder while $e = 1$ for the torus. Now, for $G = $U(1), we have $\int_G d\mu_H(g) g^n = \delta(n, 0)$. Hence, it follows immediately that the characteristic functional is non-zero if and only if $l_x = l_y = 0$. Therefore, we will focus on this case in the sequel.

Now, let us consider the partition function, Z. For the plane and the cylinder, different plaquette contributions decouple and we obtain:

$$Z = \left[\int_G d\mu_H(g) \exp(1 - \Re(g)) \right]^{N_x N_y}. \tag{IV.1}$$

For the torus, on other the hand, decoupling does not occur and we are left with

$$Z = \int \prod_\square d\mu_H(g_\square) \exp(-(1 - \Re(g_\square))) \delta\left(\prod_\square g_\square, 1_N \right). \tag{IV.2}$$

Thus, even in the Abelian, U(1) case, the Euclidean theory in two space–time dimensions has interactions! We will continue the discussion of this case in Appendix B.

Collecting these results, for the plane and the cylinder, we can now reduce the expression of $\chi(\alpha)$ to:

$$\chi(\alpha) = \prod_{I=1}^n \left[\frac{\int_G d\mu_H(g) \exp(-\beta(1 - \Re(g))) g^{k_I}}{\int_G d\mu_H(g) \exp(-\beta(1 - \Re(g)))} \right]^{|\alpha_I|} \tag{IV.3}$$

in case when $l_x = 0$ (and $\chi(\alpha) = 0$ otherwise.) We now want to take the continuum limit. The ultra-violet limit corresponds to letting lattice spacing go to zero, i.e., $\beta \to \infty$, and the infrared limit corresponds to letting the lattice size go to infinity, i.e., $L_x \to \infty$ and $L_y \to \infty$.

Let us set

$$J_n(\beta) := \int_G d\mu_H(g) \exp(-\beta(1 - \Re(g))) g^n. \tag{IV.4}$$

Now, since g is simply a complex number of modulus one it is obvious that the fraction $J_n(\beta)/J_0(\beta)$ in (IV.3) is a real number of modulus less than or equal to one. Now observe that $|\alpha_I| = g_0^2 \beta A(\alpha_I)$, where $A(\alpha_I)$ is the Euclidean area enclosed by α_I. In the limit, $\beta \to \infty$, the

integrand of both numerator and denominator become concentrated at $g = 1$, whence we have an expansion of the form $J_n/J_0 = (1 - c(1,n)/\beta)(1 + O(1/\beta^2))$, where c is positive because J_n/J_0 approaches the value 1 from below. Thus, it is easy to see that

$$\lim_{\beta \to \infty} \chi(\alpha) = \exp\left(-g_0^2 \sum_{I=1}^{n} c(1,k_I) A(\alpha_I) \right) \tag{IV.5}$$

for $l_x = 0$ and zero otherwise. We will calculate the coefficients $c(1,n)$ in section IV C. Finally, note that the infra-red limit is trivial since $\chi(\alpha)$ is independent of L_x, L_y, (assuming of course that they are large enough for the region under consideration to contain the loop).

To summarize, we can arrive at the continuum characteristic functions as follows. Given any piecewise analytic loop α' in M, we first consider a sufficiently fine and sufficiently large lattice and approximate α' by a loop α lying in the lattice. Then, we express α as a product of non-overlapping simple loops and compute the regulated characteristic function $\chi(\alpha)$ directly. Finally, we take continuum limit to arrive at the final expression (IV.5).

We will conclude this sub-section by pointing out that the Abelian case has been discussed in the literature extensively (see, e.g., Ref. 5 for an early treatment) and is included here mainly for completeness.

B. Non-Abelian case (SU(*N*))

Let us now consider the technically more difficult non-Abelian case. As indicated before, in this discussion, we will restrict ourselves to the plane and the cylinder.

For SU(N), the trace identities only enable one to express traces of products of matrices as linear combinations of traces of products of $r := N - 1$ or fewer matrices. Hence, unlike in the Abelian case, the product $T_{\alpha_1} \cdots T_{\alpha_n}$ can not be reduced to a single T_α; we can no longer confine ourselves to single loops. Fix a multi-loop — i.e., a set of r loops — $\alpha_1, \ldots, \alpha_r$ and consider its decomposition into simple, non-overlapping loops. Suppose that, in this decomposition, there are n homotopically trivial loops $\hat{\alpha}_I$ and c homotopically nontrivial loops γ_i (clearly, $c = 0$ or $c = 1$). Let $|\hat{\alpha}_I|$ be the number of plaquettes enclosed by $\hat{\alpha}_I$ and let $k_I^\pm$ and $l_i^\pm$ be the number of times that $\hat{\alpha}_I$ and γ_i occur (respectively) with positive or negative power in this decomposition. Thus, altogether, there are $b = \sum_{I=1}^{n} [k_I^+ + k_I^-] + \sum_{i=1}^{m} [l_i^+ + l_i^-]$ factors of holonomies around the $\hat{\alpha}_I, \gamma_i$ and their inverses involved in the expansion of the product $T_{\alpha_1} \cdots T_{\alpha_r}$. These may occur in arbitrary order, depending on the specific loops α_i, $i = 1, \ldots, r$.

It is then easy to see that we can now write $T_{\alpha_1} \cdots T_{\alpha_r}$ explicitly as a product of matrices representing holonomies around simple loops, with an appropriate contraction of matrix-indices:

$$N^r T_{\alpha_1} \cdots T_{\alpha_r} = \prod_{I=1}^{n} \left[\prod_{\mu=1}^{k_I^+} (h_{\hat{\alpha}_I})^{A_\mu^{I+}}_{B_\mu^{I+}} \prod_{\mu=1}^{k_I^-} (h_{\hat{\alpha}_I}^{-1})^{A_\mu^{I-}}_{B_\mu^{I-}} \right] \prod_{i=1}^{c} \left[\prod_{\nu=1}^{l_i^+} (h_{\gamma_i})^{C_\nu^{i+}}_{D_\nu^{i+}} \prod_{\nu=1}^{l_i^-} (h_{\gamma_i}^{-1})^{C_\nu^{i-}}_{D_\nu^{i-}} \right] \prod_{k=1}^{b} \delta^{E_k}_{F_{\pi(k)}}. \tag{IV.6}$$

Here, we have the following relation between indices that are being contracted:

$$(E_1, \ldots, E_b) \equiv (A_1^{1+}, \ldots, A_{k_1^+}^{1+}, A_1^{2+}, \ldots, A_{k_2^+}^{2+}, \ldots, A_1^{n+}, \ldots, A_{k_n^+}^{n+},$$

$$A_1^{1-}, \ldots, A_{k_1^-}^{1-}, A_1^{2-}, \ldots, A_{k_2^-}^{2-}, \ldots, A_1^{n-}, \ldots, A_{k_n^-}^{n-},$$

$$C_1^{1+}, \ldots, C_{l_1^+}^{1+}, C_1^{2+}, \ldots, C_{l_2^+}^{2+}, \ldots, C_1^{n+}, \ldots, C_{l_n^+}^{c+},$$

$$C_1^{1-}, ..., C_{l_1}^{1-}, C_1^{2-}, ..., C_{l_2}^{2-}, ..., C_1^{n-}, ..., C_{l_n}^{c-}), \tag{IV.7}$$

and similarly with the exchanges $E \leftrightarrow F$, $A \leftrightarrow B$, $C \leftrightarrow D$; and π is an element of the symmetric group of b elements that depends on the loops α_i and defines the specific contraction involved in $T_{\alpha_1}, \cdots T_{\alpha_r}$.

To evaluate the expectation values of this product of traces of holonomies, we need to expand out the inverses of matrices that appear in (IV.6) explicitly. This can be done easily using the fact that the matrices in question are all uni-modular. We have:

$$(h_{\hat{\alpha}_I}^{-1})_{B_\mu^{I-}}^{A_\mu^{I-}} = \frac{1}{(N-1)!} \epsilon^{A_\mu^{I-} E_{\mu,1}^I \cdots E_{\mu,N-1}^I} \epsilon_{B_\mu^{I-} F_{\mu,1}^I \cdots F_{\mu,N-1}^I} (h_{\hat{\alpha}_I})_{E_{\mu,1}^I}^{F_{\mu,1}^I} \cdots (h_{\hat{\alpha}_I})_{E_{\mu,N-1}^I}^{F_{\mu,N-1}^I}$$

$$=: \mathscr{E}_{B_\mu^{I-} F_{\mu,1}^I \cdots F_{\mu,N-1}^I}^{A_\mu^{I-} E_{\mu,1}^I \cdots E_{\mu,N-1}^I} (h_{\hat{\alpha}_I})_{E_{\mu,1}^I}^{F_{\mu,1}^I} \cdots (h_{\hat{\alpha}_I})_{E_{\mu,N-1}^I}^{F_{\mu,N-1}^I},$$

and similarly for the inverse of h_{γ_i}. Finally, if we define $n_I := k_I^+ + (N-1)k_I^-$, $c_i := l_i^+ + (N-1)l_i^-$ we can rewrite (IV.6) using a tensor-product notation as:

$$N^r T_{\alpha_i} \cdots T_{\alpha_r} = \prod_{I=1}^n (\otimes^{n_I} h_{\alpha_I})_{B_1^{I+} \cdots B_{k_I^+}^{I+} E_{1,1}^I \cdots E_{1,N-1}^I \cdots E_{k_I^-,1}^I \cdots E_{k_I^-,N-1}^I}^{A_1^{I+} \cdots A_{k_I^+}^{I+} F_{1,1}^I \cdots F_{1,N-1}^I \cdots F_{k_I^-,1}^I \cdots F_{k_I^-,N-1}^I}$$

$$\times \prod_{i=1}^c (\otimes^{c_i} h_{\gamma_i})_{D_1^{i+} \cdots D_{l_i^+}^{i+} G_{1,1}^i \cdots G_{1,N-1}^i \cdots G_{l_i^-,1}^i \cdots G_{l_i^-,N-1}^i}^{C_1^{i+} \cdots C_{l_i^+}^{i+} H_{1,1}^i \cdots H_{1,N-1}^i \cdots H_{l_i^-,1}^i \cdots H_{l_i^-,N-1}^i}$$

$$\times \prod_{k=1}^b \delta_{F_{\pi(k)}}^{E_k} \prod_{I=1}^n \prod_{\mu=1}^{k_I^-} \mathscr{E}_{B_\mu^{I-} F_{\mu,1}^I \cdots F_{\mu,N-1}^I}^{A_\mu^{I-} E_{\mu,1}^I \cdots E_{\mu,N-1}^I} \prod_{i=1}^c \prod_{\nu=1}^{l_i^-} \mathscr{E}_{D_\nu^{i-} H_{\nu,1}^i \cdots H_{\nu,N-1}^i}^{C_\nu^{i-} G_{\nu,1}^i \cdots G_{\nu,N-1}^i}. \tag{IV.8}$$

Next, let us examine the contributions from homotopically trivial loops. Choose $\eta := \hat{\alpha}_I$ for some I and consider the expression

$$(\otimes^n h_\eta)_{B_1 \cdots B_n}^{A_1 \cdots A_n}. \tag{IV.9}$$

Label the plaquette loops enclosed by η from 1 to $|\eta| := m$; thus $h_\eta = g_1 \cdots g_m$, where $g_k := h_{\square_k}$. Then the above expression becomes

$$[(g_1)_{C_{1,1}}^{A_1} (g_2)_{C_{1,2}}^{C_{1,1}} \cdots (g_m)_{B_1}^{C_{1,m-1}}] \cdots [(g_1)_{C_{n,1}}^{A_n} (g_2)_{C_{n,2}}^{C_{n,1}} \cdots (g_m)_{B_n}^{C_{n,m-1}}]$$

$$= [\otimes^n g_1]_{C_{1,1} \cdots C_{n,1}}^{A_1 \cdots A_2} [\otimes^n g_2]_{C_{1,2} \cdots C_{n,2}}^{C_{1,1} \cdots C_{n,1}} \cdots [\otimes^n g_m]_{B_1 \cdots B_n}^{C_{1,m-1} \cdots C_{n,m-1}}$$

$$= ([\otimes^n g_1][\otimes^n g_2] \cdots [\otimes^n g_m])_{B_1 \cdots B_n}^{A_1 \cdots A_n} \tag{IV.10}$$

where, in the last step we have used the product rule for tensor products of matrices.

With these explicit expressions at hand, we can now consider the functional integral which yields the Wilson loop functionals. In this evaluation, each of the n-fold tensor products in (IV.10) has to be integrated with the measure

$$d\mu(g) = d\mu_H(g) \exp(\beta/N \Re \, \mathrm{tr}(g)). \tag{IV.11}$$

To carry out this task, we will use the representation theory reviewed in Appendix A.

According to Appendix A, we have:

$$\int_G d\mu(g)\otimes^n g = \oplus_{\{m\}}\oplus_{i=1}^{f_{\{m\}}^{(n)}}\int_G d\mu(g)[p_{\{m\},i}^{(n)}\otimes^n g]$$

$$= \oplus_{\{m\}}\oplus_{i=1}^{f_{\{m\}}^{(n)}}[p_{\{m\},i}^{(n)}\otimes^n 1_N]\frac{1}{d_{\{m\}}}\int_G d\mu(g)\,\mathrm{tr}([p_{\{m\},i}^{(n)}\otimes^n g])$$

$$= \oplus_{\{m\}}[\oplus_{i=1}^{f_{\{m\}}^{(n)}}[p_{\{m\},i}^{(n)}\otimes^n 1_N]]\frac{1}{d_{\{m\}}}\int_G d\mu(g)\chi_{\{m\}}(g)$$

$$= \oplus_{\{m\}}[p_{\{m\}}^{(n)}\otimes^n 1_N]\frac{1}{d_{\{m\}}}\int_G d\mu(g)\chi_{\{m\}}(g)$$

$$= \oplus_{\{m\}}[p_{\{m\}}^{(n)}\otimes^n 1_N]J_{\{m\}}(\beta,N). \tag{IV.12}$$

Here, in the first step, we have decomposed the matrix $\otimes^n g$ into a direct sum of irreducible representations, with i labeling the orthogonal equivalent representations and m labeling the equivalence classes of inequivalent representations, and $p_{\{m\},i}^{(n)}$ are the Young symmetrizers; in the third step, we have used the fact that the trace is a class function ($\chi_{\{m\}}$ being the character of the representation $\{m\}$); and, in the last step we have simply defined

$$J_{\{m\}}(\beta,N):=\frac{1}{d_{\{m\}}}\int_G d\mu(g)\chi_{\{m\}}(g). \tag{IV.13}$$

Finally, using the orthogonality of the projectors $p_{\{m\}}^{(n)}\otimes^n 1_N$ we find that the integral over (IV.9) becomes

$$\left(\sum_{\{m\}}[p_{\{m\}}^{(n)}\otimes^n 1_N][J_{\{m\}}(\beta,N)]^{|\eta|}\right)_{B_1\cdots B_n}^{A_1\cdots A_n}. \tag{IV.14}$$

The integral over the homotopically non-trivial loops is quite similar, the main difference being that the measure there is the Haar measure and that each of these loops involves just a single integration variable. According to Appendix A we find that the integral over $\otimes^n g$ with the Haar measure is given by

$$[p_0^{(n)}\otimes^n 1_N],$$

where $p_0^{(n)}$ is the projector on the trivial representation.

Collecting these results, we can write the vacuum expectation value of $T_{\alpha_1},\ldots,T_{\alpha_r}$ as follows. Set

$$J_0(N,\beta):=\int_G d\mu(g). \tag{IV.15}$$

Then,

$$N^r \chi(\alpha_1,\ldots,\alpha_r) = \prod_{I=1}^{n} \left(\sum_{\{m\}} \left[\frac{J_{\{m\}}}{J_0} \right]^{|\hat{\alpha}_I|} [p_{\{m\}}^{(n_I)} \otimes^{n_I} 1_N] \right)_{B_1^{I+} \cdots B_{k_I^+}^{I+} E_{1,1}^{I} \cdots E_{1,N-1}^{I} \cdots E_{k_I^-,1}^{I} \cdots E_{k_I^-,N-1}^{I}}^{A_1^{I+} \cdots A_{k_I^+}^{I+} F_{1,1}^{I} \cdots F_{1,N-1}^{I} \cdots F_{k_I^-,1}^{I} \cdots F_{k_I^-,N-1}^{I}}$$

$$\times \prod_{i=1}^{m} ([p_0^{(m_i)} \otimes^{m_i} 1_N])_{D_1^{i+} \cdots D_{l_i^+}^{i+} G_{1,1}^{i} \cdots G_{1,N-1}^{i} \cdots G_{l_i^-,1}^{i} \cdots G_{l_i^-,N-1}^{i}}^{C_1^{I+} \cdots C^{i+} l_i^+ H_{1,1}^{i} \cdots H_{1,N-1}^{i} \cdots H_{l_i^-,1}^{i} \cdots H_{l_i^-,N-1}^{i}}$$

$$\times \prod_{k=1}^{b} \delta_{F_{\pi(k)}}^{E_1} \prod_{I=1}^{n} \prod_{\mu=1}^{k_I^-} \mathscr{E}_{B_\mu^{I-} F_{\mu,1}^{I} \cdots F_{\mu,N-1}^{I}}^{A_\mu^{I-} E_{\mu,1}^{I} \cdots E_{\mu,N-1}^{I}} \prod_{i=1}^{m} \prod_{\nu=1}^{l_i^-} \mathscr{E}_{D_\nu^{i,-} H_{\nu,1}^{i} \cdots H_{\nu,N-1}^{i}}^{C_\nu^{i-} G_{\nu,1}^{i} \cdots G_{\nu,N-1}^{i}}. \quad \text{(IV.16)}$$

This is the closed expression for the regulated Wilson loops. Although it seems complicated at first, its structural form is rather simple. (A more elegant derivation of (IV.16) uses the notion of a loop-network state,[9] however, since products of traces of the holonomy are more familiar to gauge theorists we have refrained from introducing the associated mathematical apparatus here.) First of all, the lattice spacing and the coupling constant enter this expression only through $J_{\{m\}}$. The rest is all an explicit contraction of indices of a product of a *finite* number of matrices. For any given group $G = \text{SU}(N)$, the matrices depend only on the decomposition of $T_{\alpha_1},\ldots,T_{\alpha_r}$ in terms of the n holonomies around the homotopically trivial, simple loops and the m holonomies around the homotopically non-trivial simple loops.

To establish the existence of the continuum limit, therefore, we only need to show that $[J_{\{m\}}(\beta)/J_0(\beta)]^{|\alpha_I|}$ converges to a finite value as $a \to 0$. Let us begin by noting that

$$|J_{\{m\}}(\beta,N)| \leq \int_G d\mu(g) \left| \frac{\chi_{\{m\}}(g)}{d_{\{m\}}^{(n)}} \right| \leq J_0(\beta,N).$$

This estimate implies that $|J_{\{m\}}/J_0|$ is always a number between 0 and 1 for finite β. Moreover, we have

$$\lim_{\beta \to \infty} \frac{J_{\{m\}}}{J_0} = \lim_{\beta \to \infty} \frac{\int_G d\mu_H \exp(-\beta(1 - 1/N \Re \, \text{tr}(g))) \dfrac{\text{tr}(\pi(g))}{\dim(\pi)}}{\displaystyle\int_G d\mu_H \exp(-\beta(1 - 1/N \Re \, \text{tr}(g)))} = 1$$

since for $\beta \to \infty$ the measure in both numerator and denominator becomes concentrated at the identity for which both integrand are equal to the number one. Therefore, we have an asymptotic expansion of the form

$$\frac{J_{\{m\}}(\beta,N)}{J_0(\beta,N)} = \left(1 - \frac{c(N,\{m\})}{\beta} \right)(1 + O(1/\beta^2)),$$

where the first order coefficient $c(N,\{m\})$ must be *non-negative* since $J_{\{m\}}/J_0$ approaches unity *from below*. Finally, observing that $|\alpha_I| = \beta g_0^2 A(\alpha_I)$, we find that the continuum limit of (IV.16) is given by replacing the $[J_{\{m\}}/J_0]^{|\alpha_I|}$ by

$$\lim_{\beta \to \infty} \left[\frac{J_{\{m\}}(\beta,N)}{J_0(\beta,N)} \right]^{|\alpha_I|} = \exp(-c(N,\{m\}) g_0^2 A(\alpha_I)). \quad \text{(IV.17)}$$

This establishes the existence of the continuum limit. To obtain the explicit formula for the Wilson loops, it only remains to evaluate the constants $c(N,\{m\})$. We will carry out this task in the next sub-section.

We will conclude this sub-section with a few remarks. (Some of these observations have been made in the context of other approaches but are included here for completeness.)

(i) The explicit expression of the Wilson loop functionals (or the characteristic functional for the Yang–Mills measure on $\overline{\mathscr{A}/\mathscr{G}}$) is rather complicated. Note however that the computation only involves complicated traces and can be performed by algebraic manipulation programs very quickly. Furthermore, some of the qualitative features can be easily read-out. Note first that if we have a single, simple loop α_0, the matrix factors in (IV.16) disappear and the expectation value collapses to simply:

$$\langle T_{\alpha_0} \rangle \equiv \chi(\alpha_0) = e^{-cg_0^2 A(\alpha_0)}, \qquad (IV.18)$$

where $2c$ is the value of the first SU(N) Casimir on its fundamental representation (see the next subsection) and where $A(\alpha_0)$ is the Euclidean area enclosed by the loop α_0. Thus, the area law — generally taken to be the signal of confinement — holds. Note that the loop does not have to be large; the expression is exact. Finally, note from section IV A that this law holds also for the *Abelian* theory. Thus, the continuum limit of the lattice U(1) theory provides us the confined phase of the theory which is different from the phase described by the standard Fock representation.

(ii) More generally, if one restricts oneself to *non-overlapping* loops $\alpha_1, \ldots, \alpha_n$, our closed expression (IV.16) yields

$$\langle \pi_1(\alpha_1) \otimes \cdots \otimes \pi_n(h_{\alpha_n}) \rangle = \langle \pi_1(h_{\alpha_1}) \rangle \otimes \cdots \otimes \langle \pi_n(h_{\alpha_n}) \rangle \qquad (IV.19)$$

with $\langle \pi(h_\alpha) \rangle = \pi(1_N) e^{-c(N,\pi)g_0^2 A(\alpha)}$ of (IV.17), where as before π is the irreducible representation. This result is agreement with the results obtained by Bralic,[5] Gross *et al.*[4] and Klimek *et al.* and Kazakov.[28] However, even for this special case, our method of arriving at the result is different. As explained in the Introduction, we do not break gauge invariance to pass to a ''kinematically linear'' case nor do we use stochastic differential equations.

(iii) Note that, as in the Abelian theory, the infra-red limit is trivial since the continuum expression of the Wilson loop functionals does not depend on L_x or L_y at all (provided of course the lattice is chosen large enough to encompass the given r loops).

(iv) It is interesting to note that we did not have to renormalize the bare coupling constant g_0 in the process of taking the continuum limit. This is a peculiarity of two dimensions. Indeed, in higher dimensions, the bare coupling constant does not have the correct physical dimensions to allow for an area law which suggests that renormalization would be essential.

(v) In the classical theory in higher dimensions, the Yang–Mills action depends on the space-time metric and is thus invariant only under the action of the finite dimensional isometry group of the underlying space–time [the Poincaré (respectively, Euclidean) group, if the space–time is globally Minkowskian (Euclidean)]. In two space–time dimensions, on the other hand, one needs only an area element to write the Yang–Mills action. Thus, the symmetry group is considerably enlarged; it is the *infinite* dimensional group of area preserving diffeomorphisms. A natural question is whether the Wilson loop functionals are also invariant under this larger group. Our explicit expression makes it obvious that it is. Thus, the infinite-dimensional symmetry is carried over in-tact to the quantum theory. This property is not obvious in many other approaches which use gauge-fixing to endow $\mathscr{A}/\mathscr{G}$ a vector space structure and then employ the standard (space–time metric dependent) Gaussian measures in the intermediate steps. In these approaches, special and somewhat elaborate calculations are needed to verify invariance under all area preserving diffeomorphisms.

(vi) As one can explicitly check, if one performs the $N \to \infty$ limit of our continuum theory on the plane or the cylinder using the above expression then one does not encounter a phase transition. This seems to contradict certain results[29] obtained in a two-dimensional finite, planar lattice theory. However, there is no contradiction because in two dimensions on the plane or the cylinder the appearance of the $N \to \infty$ phase transition is a lattice artefact. In Ref. 17 the authors observe a third order phase transition in the limit $N \to \infty$ also in the continuum. However, those authors consider the case that the two-dimensional manifold is a sphere rather than a plane or a cylinder so that again there is no contradiction.

C. Determination of the coefficients $c(N, \{m\})$

The main idea behind the calculation is the following; Since for $\beta \to \infty$, the integrand of $J_{\{m\}}(\beta)$ is concentrated at the identity, it is sufficient to calculate the integrand in Eq (IV.13) (defining $J_{\{m\}}$) in a neighborhood of the identity.

To that effect, write $g = e^A$, where $A = t^I \tau_I \in L(G)$ is in the Lie algebra of G and t^I are real parameters in a neighborhood of zero. We thus have upon inserting $g = 1_N + A + \frac{1}{2} A^2 + o(A^3)$

$$1 - \frac{1}{N} \Re \operatorname{tr}(g) = -\frac{1}{2N} \operatorname{tr}(A^2) + o(A^3) = \frac{1}{2} \sum_{I=1}^{\dim(G)} (t^I)^2 + o(A^3), \tag{IV.20}$$

where the term of first order in A vanishes because it is either purely imaginary [the Abelian sub-ideal of $L(G)$] or trace-free [the semi-simple sub-ideal of $L(G)$] and where we have used the normalization $\operatorname{tr}(\tau_I \tau_J) = -N \delta_{IJ}$. Similarly, we have an expansion for the $\{m\}$th irreducible representation of G given by $\pi_{\{m\}}(g) = 1_{\{m\}} + X + \frac{1}{2} X^2 + o(X^3)$, where $X = t^I X_I$ is the representation of the Lie algebra element A in the $\{m\}$th irreducible representation. Then we have

$$\chi_{\{m\}}(g) = d_{\{m\}} + t^I \operatorname{tr}(X_I) + \frac{1}{2} t^I t^J \operatorname{tr}(X_I X_J) + o(X^3). \tag{IV.21}$$

Now, according to the Baker–Campbell–Hausdorff formula[30] we have:

$$e^{t^I \tau_I} e^{s^I \tau_I} = e^{r^I(s,t) \tau_I}, \quad \text{where} \quad r^I(s,t) = s^I + t^I - \frac{1}{2} f^I_{JK} s^J t^K + o(s^2, t^2, s^3, s^2 t, s t^2, t^3) \tag{IV.22}$$

and where f^I_{JK} are the structure constants of the semi-simple sub-ideal of $L(G)$ which therefore are completely skew. Finally, the Haar measure can be written[30]

$$d\mu_H(e^{t^I \tau_I}) = \frac{d^{\dim(G)} t}{\det \left(\dfrac{\partial r^I(s,t)}{\partial s^J} \right)_{s=0}} = \frac{d^{\dim(G)} t}{1 + o(t^2)} \tag{IV.23}$$

since $\det(\partial r / \partial s)_{s=0} = \det(1 + \frac{1}{2} t^I R_I + o(t^2)) = 1 + \frac{1}{2} \operatorname{tr}(t^I R_I) + o(t^2) = 1 + o(t^2)$, where $(R_I)^J_K = f^J_{IK}$ is the Ith basis vector of the semi-simple sub-ideal of $L(G)$ in the adjoint representation which is trace-free.

We are now ready to carry out the required estimate. There exists a subset $U_0 \subset \mathbb{R}^{\dim(G)}$ which is in one-to-one correspondence with G via the exponential map. Let U be the closure of U_0 in $\mathbb{R}^{\dim(G)}$. The set U is compact in $\mathbb{R}^{\dim(G)}$ because G is compact and so the set U_0 must be bounded. Furthermore, since the group under consideration has only a finite number of connected components (namely, one), there are also only a finite number of corresponding connected components of U_0 and therefore the set $U - U_0$ has at most dimension $\dim(G) - 1$. It follows that $U - U_0$ has Lebesgue measure zero, that is, we can replace the integral over U_0 with respect to $d^{\dim(G)} t$ by an integral over U. For instance, for U(1) the set U is just given by the interval $[-\pi, \pi]$ while U_0 could be chosen as $[-\pi, \pi)$. Likewise, for SU(2) the set U is the set of points

$t_1^2 + t_2^2 + t_3^2 \leqslant \pi$ while U_0 is the set of points $t_1^2 + t_2^2 + t_3^2 < \pi$ plus one arbitrary additional point of radius π corresponding to the element -1_2. Inserting (IV.19), (IV.20) and (IV.22) into (IV.13) we can therefore write an expansion in $1/\sqrt{\beta}$

$$
\begin{aligned}
d_{\{m\}}[J_{\{m\}}(\beta) - J_0(\beta)] &= \int_U \frac{d^{\dim(G)}t}{1 + o(t^2)} \, \exp\!\left(-\beta \frac{1}{2} \sum_{I=1}^{\dim(G)} (t^I)^2 + \beta o(t^3) \right) \\
&\quad \times \left[t^I \, \mathrm{tr}(X_I) + \frac{1}{2} t^I t^J \, \mathrm{tr}(X_I X_J) + o(t^3) \right] \\
&= \frac{1}{\beta^{\dim(G)/2+1}} \int_{\sqrt{\beta}U} \frac{d^{\dim(G)}t}{1 + o(t^2/\beta)} \, \exp\!\left(-\frac{1}{2} \sum_{I=1}^{\dim(G)} (t^I)^2 + o(t^3/\sqrt{\beta}) \right) \\
&\quad \times \left[\sqrt{\beta} t^I \, \mathrm{tr}(X_I) + \frac{1}{2} t^I t^J \, \mathrm{tr}(X_I X_J) + o(t^3/\sqrt{\beta}) \right] \\
&= \frac{1}{\beta^{\dim(G)/2+1}} \int_{\mathbb{R}^{\dim(G)}} d^{\dim(G)}t \, \exp\!\left(-\frac{1}{2} \sum_{I=1}^{\dim(G)} (t^I)^2 \right) \\
&\quad \times \left[\sqrt{\beta} t^I \, \mathrm{tr}(X_I) + \frac{1}{2} t^I t^J \, \mathrm{tr}(X_I X_J) + o(t^3/\sqrt{\beta}) \right],
\end{aligned}
\tag{IV.24}
$$

where in the last step the expansion of the scaled domain $\sqrt{\beta}U$, U a compact subset of $\mathbb{R}^{\dim(G)}$ to all of $\mathbb{R}^{\dim(G)}$ also is correct up to a further order in $1/\sqrt{\beta}$. Now the terms of odd order in t vanish due to the symmetry of the exponential under reflection. Therefore, we have:

$$
d_{\{m\}}[J_{\{m\}}(\beta) - J_0(\beta)] = \frac{1}{\beta} J_0(\beta) \frac{1}{2} \mathrm{tr}\!\left(\sum_{I=1}^{\dim(G)} (X^I)^2 \right) + o(1/\beta^2).
\tag{IV.25}
$$

But $\Sigma_I (X_I)^2 = -\lambda_{\{m\}} 1_{\{m\}}$ is the Casimir invariant and $\lambda_{\{m\}}$ is its eigenvalue. Therefore we arrive finally at

$$
c(N, \{m\}) = \tfrac{1}{2} \lambda_{\{m\}}.
\tag{IV.26}
$$

It is well-known[31] that the Laplace–Beltrami operator $-\Delta$ has eigenvalues $\lambda_{\{m\}}$ on its complete system of conjugation invariant eigenfunctions $\chi_{\{m\}}(g)$. These functions are parametrized by r discrete quantum numbers, according to the rank of G.

V. THE HAMILTONIAN FORMALISM

In this section, we will recall the standard Hamiltonian formulation of Lorentzian Yang–Mills theory in $1+1$ dimensions. (For details, see, e.g., Refs. 32–34). Here we will only consider topologies $M = \mathbb{R}^2$ and $M = S^1 \times \mathbb{R}$ since the Lorentzian metric, obtained by analytic continuation, on the torus $S^1 \times S^1$ has closed time-like curves. This discussion will be used in section V C to show the equivalence of our Euclidean framework with the standard Hamiltonian description.

The canonical form of the Yang–Mills actions is given by

$$
S = \int_R dt \int_\Sigma dx \left[\dot{A}_I E^I - \left[-\Lambda^I \mathscr{G}_I + \frac{g_0^2}{2} E^I E^I \right] \right],
\tag{V.1}
$$

where $\Sigma = \mathbb{R}$ or S^1 and a dot (prime) denotes a derivative with respect to t (x). Here $A = A_x$ is the the x-component of the G connection and $E = (1/g_0^2)(\partial_t A_x - \partial_x A_t + [A_t, A_x])$ is its electric field. The indices I, J, K run $1, \ldots, \dim(G)$ and are raised and lowered with respect to the Cartan Killing metric. Note that time component $A_t^I = \Lambda^I$ of the connection acts as a Lagrange multiplier, enforcing the Gauss constraint

$$\mathscr{G}_I = E_I{}' + [A, E]_I. \tag{V.2}$$

Because the magnetic fields vanish in one spatial dimension, the Hamiltonian takes the form

$$H = \int_\Sigma dx\, \frac{g_0^2}{2} E^I E^I. \tag{V.3}$$

However, multiplying the Gauss constraint by E^I yields

$$\tfrac{1}{2}(E^I E^I)' = 0$$

so that the Hamiltonian density must be a constant. Thus, the energy on the plane is infinite unless that constant is zero. This enforces the new first class constraints $E^I = 0$. The motions generated by these constraints are transitive on the whole configuration space of the A_I and so A_I is identified with the trivial connection $A_I = 0$. The reduced phase space for $M = \mathbb{R}^2$ is therefore zero-dimensional, it consists only of one point, $(0,0)$, say.

Remark: A more interesting theory results if we weaken the boundary conditions to allow non-zero electric fields at infinity. For definiteness, let us consider the SU(2) theory and define the phase space as follows: (A_I, E^I) belong to the phase space if $A_I = O(1/x^2)$ and $E^I \mapsto E_o \hat{v}^I$ as $x \mapsto \pm\infty$, where E_o is an arbitrary constant and v^I is a fixed internal vector. It is easy to check that the symplectic structure is well-defined on this phase space. Physically, the boundary conditions ensure that we have ''an external electric field.'' (The previous arguments do imply that the total Hamiltonian of the system is infinite but the energy per unit length is finite.) The Gauss law again generates gauge transformations which are asymptotically identity. We can partially fix this gauge freedom by demanding that the electric field be everywhere parallel to v^I. Then the Gauss constraint itself implies that $E^I = E_o v^I$ everywhere and that A_I is also parallel to v_I. The remaining gauge freedom can be exhausted by bringing A_I to a standard form: $A_I = A_o f(x) v^I$, where $f(x)$ is a fixed function and the value of the constant A_o is determined by the holonomy of the given connection $A_I(x)$. This exhausts the gauge freedom and solves the Gauss law. The true degrees thus captured in the pairs (A_o, E_o); the reduced phase space is topologically $\mathbb{R}^2$. [For the SU(N) theory, it is $\mathbb{R}^{2r}$.] We will not treat these cases in any detail here, however, because the modifications needed to incorporate these ''external fields'' in the Euclidean description is beyond the scope of this work (as well as of other mathematical physics treatments that we are aware of).

On the cylinder, the theory is analogous to the more general case discussed above but the Hamiltonian is now finite. It is given by

$$H = \frac{g_0^2 L_x}{2}(E^I E^I). \tag{V.4}$$

By a gauge transformation,[34] we may take A, E to be constant. By means of a constant gauge transformation we achieve that A lies in a Cartan subalgebra. Since in that gauge the Gauss constraint implies that A, E commute, it follows that there is a gauge in which A, E both lie in a

Cartan subalgebra. Let r be the rank of $L(G)$; then the maximal Cartan subalgebra has dimension r and the reduced phase space has dimension $2r$. The reduced phase space is then the quotient[34] of $\mathbb{R}^{2r}$ by a discrete set of residual gauge transformations.

In the quantum theory on a cylinder, the Hamiltonian becomes the Laplace Beltrami operator on the Cartan subgroup G_C[31]

$$H = -\frac{g_0^2 L_x}{2}\Delta \tag{V.5}$$

and physical states correspond to conjugation invariant functions on G. The corresponding inner product is the L^2 inner product given by the Haar measure on G_C. As a result, the characters $\chi_{\{m\}}(A)$ with $\{m\} = \{m_1, \ldots, m_r\}$ with $m_1 \geq m_2 \geq \cdots \geq m_r \geq 0$ provide a complete set of eigenstates of H (with eigenvalues $g_0^2 L_x \lambda_{\{m\}}/2$). For comparison with the classical theory, recall that the characters $\chi_{\{m\}}$ depend only on the Cartan subgroup of G.

VI. AXIOMATIC FRAMEWORK AND RELATION TO THE HAMILTONIAN THEORY

In scalar field theories, the Osterwalder–Schrader axiomatic framework provides a compact formulation of what is often referred to as ''the main problem.'' Consequently, the framework plays a central role in constructive quantum field theory. However, as mentioned in the Introduction, this framework is geared to ''kinematically linear'' theories because a basic premise of the axioms is that the space of paths is a vector space, generally taken to be the space $\mathscr{S}'$ of tempered distributions. In this section, we will use the material presented in sections II and IV to suggest a possible generalization of the Osterwalder–Schrader framework to gauge theories, using for the space of physical paths the non-linear space $\overline{\mathscr{A}/\mathscr{G}}$.

The section will be divided into three parts. In the first, we briefly review the aspects of the Osterwalder–Schrader framework that are relevant for our discussion. In the second, we propose an extension of the key axioms and verify that they are satisfied by the continuum SU(N) Yang–Mills theories. In the third part we show that the axioms suffice to demonstrate the equivalence between the Euclidean and the Hamiltonian frameworks.

A. Kinematically linear theories

As mentioned in the Introduction, the basic idea of the Euclidean constructive quantum field theory[6] is to *define* a quantum field theory through the measure μ on the space of paths Φ—the rigorous analog of ''exp$-S(\Phi)\mathscr{D}\Phi$.'' In the Osterwalder–Schrader framework, the space of paths is taken to be the space $\mathscr{S}'$ of tempered distributions on the Euclidean space–time $\mathbb{R}^d$, and conditions on permissible measures μ on $\mathscr{S}'$ are formulated as axioms on their Fourier transforms $\chi(f)$, defined via

$$\chi(f) := \langle\exp(i\overline{\Phi}[f])\rangle := \int_{\mathscr{S}'} d\mu(\overline{\Phi})\exp(i\overline{\Phi}[f]). \tag{VI.1}$$

Here f are test functions in the Schwartz space $\mathscr{S}$, the over-bar is used to emphasize that the fields are distributional and $\overline{\Phi}[f] = \int_{\mathbb{R}^d} d^d x \overline{\Phi}(x) f(x)$ denotes the canonical pairing between distributions and test functions. The generating functional $\chi(f)$ determines the measure completely. Furthermore, since $\mathscr{S}$ is a nuclear space, Minlos' theorem[25] ensures that if we begin with *any* continuous, positive linear functional χ on $\mathscr{S}$, there exists a regular measure μ on $\mathscr{S}'$ such that (VI.1) holds.

In the Osterwalder–Schrader framework, then, *a quantum field theory is a normalized measure μ on $\mathscr{S}'$, or, equivalently, a continuous, positive linear functional χ on $\mathscr{S}$ satisfying the following axioms:*

(OS-I) *Analyticity.* This assumption ensures that the measure μ has an appropriate ''fall-off.'' It requires that $\chi(\Sigma_{i=1}^{n} z_i f_i)$ is entire analytic on $\mathbb{C}^n$ for every finite dimensional subspace spanned by the linearly independent vectors $f_i \in \mathscr{S}$.

(OS-II) *Regularity.* These are technical assumptions which, roughly speaking, allow one to construct Euclidean field operators such that its Schwinger distributions

$$S(x_1, \ldots, x_n) := \langle \overline{\Phi}(x_1), \ldots, \overline{\Phi}(x_n) \rangle$$

are *tempered* — rather than less well-behaved — distributions. We will not display them here.

(OS-III) *Euclidean invariance.* This condition ensures Poincaré invariance of the Wick-rotated theory. If gf is the image of a test function f under the action of an element g of the *full* Euclidean group E in d dimensions then, one requires:

$$\chi(gf) = \chi(f).$$

Here the test functions are considered as scalars, that is $(gf)(x) := f(gx)$.

(OS-IV) *Reflection positivity.* This is perhaps the key axiom because it enables one to reformulate the theory in terms of more familiar concepts by providing a notion of time, a Hilbert space, and a Hamiltonian. The precise condition can be formulated as follows. Choose an arbitrary hyper-plane in $\mathbb{R}^d$ which we will call the time zero plane. Consider the linear space, denoted V, generated by finite linear combinations of the following functions on $\mathscr{S}'$

$$\Psi_{\{z_i\},\{f_i\}} : \mathscr{S}' \to \mathbb{C}; \quad \overline{\Phi} \to \sum_{i=1}^{n} z_i \exp(\overline{\Phi}[f_i]),$$

where $z_i \in C$, $f_i \in \mathscr{S}$ with support only in the ''positive time'' part of the space–time $(\mathrm{supp}(f_i) = \{x = (x^0, \vec{x}) \in \mathbb{R}^d; \ x^0 > 0\})$. Next, let $\Theta(x^0, \vec{x}) = (-x^0, \vec{x})$ denote the time reflection operator $(\Theta \in E)$. Then, one requires that

$$(\Psi, \Xi) := \langle \Theta \Psi, \Xi \rangle := \int_{\mathscr{S}'} d\mu(\overline{\Phi})(\Theta \Psi[\overline{\Phi}])^{\star} \Xi[\overline{\Phi}] \geq 0 . \tag{VI.2}$$

(OS-V) *Clustering.* This axiom ensures uniqueness of the vacuum. It requires that the measure has the cluster property, that is,

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t ds \langle \Psi T(s) \Xi \rangle = \langle \Psi \rangle \langle \Xi \rangle$$

for all Ψ, Ξ in a dense subspace of $L_2(\mathscr{S}', d\mu)$. Here $T(s)$ is a representation on $L_2(\mathscr{S}', d\mu)$ of the one-parameter semi-group of time translations defined by $\hat{T}(s) \exp(\overline{i\Phi(f)}) := \exp(i\overline{\Phi}(T(s)f)$ and extended by linearity and $(T(s)f)(x^0, \vec{x}) := f(x^0 + s, \vec{x})$ for all $f \in \mathscr{S}$.

With these axioms at hand, one can construct a Hilbert space $\mathscr{H}$ of quantum states, a Hamiltonian H and a unique vacuum vector Ω (annihilated by H) as follows:

(1) Consider the *null space* $\mathscr{N}$ of norm zero vectors in V with respect to the bilinear form $(,)$ introduced in (VI.2) and Cauchy-complete the quotient $V/\mathscr{N}$. Then $\mathscr{H} := \overline{V/\mathscr{N}}$ with scalar product $(,)$.

(2) The most important theorem now is that, given a probability measure μ satisfying reflection positivity and Euclidean invariance, the time translation operator $T(s)$ acting on V factors through the quotient construction referred to in (1), that is, it leaves the null space $\mathscr{N}$ invariant. This means that we can represent it on $\mathscr{H}$ and standard Hilbert space techniques now ensure that $T(t)$ has a positive self-adjoint generator H such that $T(t) = \exp(-tH)$ [note that (due to

Euclidean invariance) $T(t)$ is unitary with respect to $\langle , \rangle$ but symmetric with respect to (,) due to the additional time reflection involved; this shows that $T(t)$ provides a symmetric contraction semi-group].

(3) The vacuum state turns out to be just the projection to $\mathscr{H}$ of the function 1 on $\mathscr{S}'$.

B. A proposal for gauge theories

The discussion of section II suggests that, in certain gauge theories, it is natural to use $\overline{\mathscr{A}/\mathscr{G}}$ as the space of physical paths. Thus, we are led to seek an extension of the Osterwalder–Schrader framework in which the linear space $\mathscr{S}'$ is replaced by the non-linear space $\overline{\mathscr{A}/\mathscr{G}}$. At first this goal seems very difficult to reach because the standard framework uses the underlying linearity in almost every step. However, we will see that one can exploit the ''non-linear duality'' between connections and loops — or, more precisely, between $\overline{\mathscr{A}/\mathscr{G}}$ and the hoop group $\mathscr{HG}$ — very effectively to extend those features of the standard framework which are essential to the proof of equivalence between the Euclidean and the Hamiltonian frameworks.

Let us consider a gauge theory in d Euclidean space–time dimensions with a compact Lie group G as the structure group. The proposal is to use $\overline{\mathscr{A}/\mathscr{G}}$ as the space of Euclidean paths. (Even though we are now working in an arbitrary dimension and with more general structure groups, this space can be again constructed using any one of the three methods discussed in section II.) Since $\overline{\mathscr{A}/\mathscr{G}}$ is compact, it admits normalized, regular Borel measures. Furthermore, the Riesz-Markov theorem (together with the Gel'fand theory) ensures[10] that each of these measures μ is completely determined by the ''characteristic functional'' $\chi(\alpha_1, \ldots, \alpha_n)$, defined by:

$$\chi(\alpha_1, \ldots, \alpha_n) := \int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu \ \check{T}_{\alpha_1} \cdots \check{T}_{\alpha_n}, \tag{IV.3}$$

where, $\check{T}_{\alpha_k}$ denotes Gel'fand transform of T_{α_k}, the trace of the holonomy around the closed loop α_k. (There is also a theorem[35] that ensures the converse, i.e., which states that given a functional χ of multi-loops satisfying certain conditions, there exists a regular measure μ on $\overline{\mathscr{A}/\mathscr{G}}$ such that χ can be reconstructed via (VI.3). However, since one has to introduce more technical machinery to state this theorem properly and since this converse is not logically necessary for the constructions that follow, we will not discuss it here.) Comparing (VI.3) and (VI.1), we see that $\overline{\mathscr{A}/\mathscr{G}}$ now plays the role of $\mathscr{S}'$ and multi-loops, the role of test functions, and traces of holonomies, the role of $\exp i\hat{\Phi}(f)$. Thus, we have extended the Fourier transform (VI.1) to a non-linear space by exploiting the fact that the loops and connections can be regarded as ''dual objects'' in the expression of the trace of the holonomy. Our strategy now is to introduce a set of axioms on measures μ through their characteristic functionals χ.

Let us begin with an observation. The discussion of the previous section brings out the fact that while all five axioms are needed to ensure that the resulting theory is complete and free of pathologies, it is the last three axioms — the Euclidean invariance and the reflection positivity — that play the central role in the reconstruction of the Hamiltonian theory. We will therefore begin with these axioms.

A quantum gauge field theory is a probability measure μ on $\overline{\mathscr{A}/\mathscr{G}}$ satisfying the following axioms:

(I) *Euclidean invariance.* μ is invariant under the full Euclidean group if the space–time topology is $\mathbb{R}^d$, and under the full isometry group of the flat Euclidean metric in more general context. In terms of characteristic function χ, we thus have:

$$\chi(g\alpha) = \chi(\alpha), \tag{IV.4}$$

where α stands for a generic multi-loop $(\alpha_1, ..., \alpha_n)$ and $g\alpha$ denotes the image of α under the action of an isometry g.

(II) *Reflection positivity.* Choose, as before, an arbitrary ''hyper-plane'' and regard it as the time-zero slice. Consider the linear space V generated by finite linear combinations of functionals on $\overline{\mathscr{A}/\mathscr{G}}$ of the form

$$\Psi_{\{z_i\},\{\alpha_{Ii}\}} : \overline{\mathscr{A}/\mathscr{G}} \to \mathbb{C}; \quad \bar{A} \to \sum_{i=1}^{n} z_i \prod_{I=1}^{r} \check{T}_{\alpha_{Ii}}(\bar{A}),$$

where the loops α_{Ii} have support in the positive half space. Then we must have:

$$(\Psi, \Xi) := \langle \Theta\Psi, \Xi \rangle := \int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu(\bar{A})(\Theta\Psi[\bar{A}])^{\star} \Xi[\bar{A}] \geqslant 0 , \tag{VI.5}$$

where, as before Θ is the time-reflection operator.

(III) *Clustering.* The requirement is the same in formulae as for the kinematically linear field theories, namely,

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t ds \langle \Psi T(s) \Xi \rangle = \langle \Psi \rangle \langle \Xi \rangle$$

for all Ψ, Ξ in a dense subspace of $L_2(\overline{\mathscr{A}/\mathscr{G}}, d\mu)$. Here $T(s)T_{\alpha_1} \cdots T_{\alpha_r} = T_{T(s)\alpha_1} \cdots T_{T(s)\alpha_r}$, where $(T(s)\alpha)^0(\tau) := \alpha^0(\tau) + s, (T(s)\vec{\alpha})(\tau) := \vec{\alpha}(\tau)$ and τ is a parameter along the loop.

We will see in the next section that these axioms suffice to reconstruct the Hamiltonian theory. However, this set of axioms is clearly incomplete (see e.g. reference 7). We will now indicate how one might impose additional conditions and point out some subtleties.

Let us begin with the analyticity axiom of Osterwalder and Schrader. In that case, we could take complex linear combinations $\Sigma z_i f_i$ because the space $\mathscr{S}$ of test functions is a vector space. In the present case, we can only compose loops (or, more precisely, hoops) to obtain

$$\alpha = \alpha_1^{n_1} \circ \cdots \circ \alpha_n^{n_n}, \quad i = 1 \; \cdots \; r,$$

with integer winding numbers n_j, and, more generally, a full subgroup of the hoop group generated by a finite number of independent hoops (the notion of ''strong independence,''[11] of hoops being the substitute for ''linear independence'' of test functions f_i.) One could also include complex winding numbers and this may lead us to the notion of ''extended loops.''[36] In any case, it may be natural to require that

$$\chi(\{\alpha_i\})$$

be ''in some sense analytic'' in the winding numbers n_{ij} (we will leave a more precise formulation of this notion for future work). Recall, however, that in the Osterwalder–Schrader framework, the analyticity axiom is needed to ensure the existence of Schwinger functions. In the present case, on the other hand, since the analogs $\langle A(x_1), ..., A(x_n) \rangle$ of the Schwinger functions fail to be gauge invariant, from our perspective, it is unnatural to require that they be well-behaved in the quantum theory. So, at this stage of our understanding, the *raison d'etre* of the analyticity condition is not as compelling in our framework. Therefore, a definitive formulation of this axiom must await further development of the framework.

The situation with the Regularity axiom is similar. In the Osterwalder Schrader framework, it prescribes certain bounds on the characteristic function $\chi(f)$ which are needed to ensure that the Schwinger functions can be continued analytically to obtain the Wightman functions in the

Lorentzian regime. In the present context, neither the Schwinger nor the Wightman functions are gauge invariant. Nonetheless, suitable regularity conditions *are* needed to ensure that the Lorentzian *Wilson loops* are well-defined. The precise form of these conditions will become clear only after the issue of analytic continuation of Wilson loops is explored in greater detail.

Finally, the space $\overline{\mathscr{A}/\mathscr{G}}$ is very large: In a well-defined sense, it serves as the "universal home" for measures in theories in which the traces of holonomies are well-defined operators.[8] From general considerations, one would expect that the measures that come from physically interesting gauge theories should have a much smaller support (provided, of course, that traces of holonomies are measurable functions). A further investigation of this issue would suggest additional restrictions on the characteristic functions.

To conclude this section, let us consider the key question that any set of axioms must face: Are they consistent? That is, do they admit *non-trivial* examples? Fortunately, results in section IV immediately imply that the answer is in the affirmative. To see this, let us take M to be either a 2-plane or a 2-cylinder and the structure group to be SU(N) or U(1). The characteristic functional is then given by (IV.16). Let us begin with Euclidean invariance. Since the characteristic functionals depend only on the areas of the various loops involved, they are invariant under all area preserving diffeomorphisms and, in particular, under the isometry groups of the underlying space–times. Reflection positivity is also satisfied because, as we will see in the next sub-section, after dividing by $\mathscr{N}$ we obtain a scalar product which is positive definite. Furthermore, since the measure is non-interacting, clustering is immediate (see next subsection). Finally, we can also test if the "obvious" restrictions of analyticity and regularity are met. By inspection, the characteristic functionals (IV.16) are formally analytic in $k_I^{\pm}$ and $l_i^{\pm}$. Since the winding numbers n_j are linear combinations of these, the generating functions are formally analytic in the winding numbers as well. Finally, the generating functionals are bounded (by 1).

C. Reconstruction of the Hamiltonian theory

Let us now construct a Hilbert space, a Hamiltonian and a vacuum via the Osterwalder Schrader algorithm[6] and verify that, for cases treated in sections IV, this description is equivalent to the one obtained directly using Hamiltonian methods in section V. Since this algorithm uses, in essence, only reflection positivity, it is directly applicable to our formulation of gauge theories.

The first step is to construct the null space $\mathscr{N}$ in V. Let us fix one of the Ψ's considered in axiom (II). Then we have

$$(\Psi,\Psi)=\sum_{i,j=1}^{n} z_i^{\star} z_j \int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu(\overline{A}) \prod_{I,J=1}^{r} (\check{T}_{\Theta(\alpha_{Ii})^{-1}}[\overline{A}])(\check{T}_{\alpha_{Jj}}[\overline{A}]),$$

where μ is the physical measure obtained by taking the continuum limit of (IV.16), and where we we have used the fact that, since G is unitary, $(\check{T}_{\alpha})^{\star}=\check{T}_{\alpha^{-1}}$, where $\star$ denotes complex conjugation.

We now need to express this equation in terms of χ. Let us begin by considering the decomposition of a multi-loop $\{\alpha_1,...,\alpha_s\}$, $s\leq r$. In this decomposition, it is convenient to separate the homotopically trivial loops from the non-trivial ones. In the case $M=\mathbb{R}\times\mathbb{R}$, there is no homotopically non-trivial loop. On the cylinder we can choose the horizontal loop γ at $t=0$ as the fiducial non-trivial loop and write every homotopically non-trivial loop η occurring in the multi-loop $\{\alpha_1,...,\alpha_s\}$ as $\eta=[\eta\circ\gamma^{-1}]\circ\gamma$, where the loop in brackets is homotopically trivial. The result will be a multi-loop $\widetilde{\alpha}_1,...,\widetilde{\alpha}_{\tilde{s}}$ whose homotopically trivial contribution comes *only* from γ. Finally write $\Pi_{I=1}^{s} T_{\alpha_I}$ as a linear combination of terms of the form [as in (IV.6)]

$$\mathrm{tr}(g_{\{m\}}(\hat{\alpha}_1,...,\hat{\alpha}_{\tilde{s}})\pi_{\{m\}}(h_{\gamma})), \tag{VI.6}$$

where $\pi_{\{m\}}$ is the $\{m\}$th irreducible representation of G and $g_{\{m\}}$ is some matrix which depends only on the homotopically trivial loops $\hat{\alpha}_i$ and which is projected from both sides by $\pi_{\{m\}}(1_N)$, that is, $g_{\{m\}}\pi_{\{m\}}(1_N) = \pi_{\{m\}}(1_N)g_{\{m\}}$. Loops $\widehat{\alpha}_i$ arise from $\widetilde{\alpha}_i$ by taking the simple loop decomposition of $\widetilde{\alpha}_i$ as in (IV.6) and taking out the γ's and its inverses. Since every multi-loop functional can be so expanded, it is sufficient to consider the scalar product among these functionals which we will now write as

$$F_{\{m\}}(\beta) := F_{\{m\}}(\beta_1, \ldots, \beta_s) := \mathrm{tr}(g_{\{m\}}(\beta_1, \ldots, \beta_s)\pi_{\{m\}}(h_\gamma)), \qquad (VI.7)$$

where β_i are homotopically trivial and enclose surfaces in the positive half-space. Note that $\Theta F_{\{m\}}(\beta) = F_{\{m\}}(\Theta\beta)$ since $\Theta\gamma = \gamma$. We can therefore alternatively write Ψ in the form

$$\Psi = \sum_{\{m\}} z_{\{m\}}F_{\{m\}}(\beta_{\{m\}}). \qquad (VI.8)$$

Now, using the formula[37]

$$\int_G d\mu_H(g)\bar{\pi}_{AB}(g) \otimes \pi'_{CD}(g) = \frac{\delta_{\pi,\pi'}}{d_\pi}\pi_{AC}(1)\pi_{BD}(1) \qquad (VI.9)$$

we find

$$\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\, \bar{F}_{\{m\}}(\Theta\beta_{\{m\}})F_{\{m'\}}(\beta_{\{m'\}}) = \frac{\delta_{\{m\},\{m'\}}}{d_{\{m\}}} \int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\, \bar{g}_{\{m\}}(\Theta\beta_{\{m\}})_{AB}g_{\{m\}}(\beta_{\{m\}})_{BA}$$

$$= \frac{\delta_{\{m\},\{m'\}}}{d_{\{m\}}}\,\mathrm{tr}\left(\left[\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\, \bar{g}_{\{m\}}(\beta_{\{m\}})\right]\left[\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\, g_{\{m\}}(\beta_{\{m\}})\right]\right)$$

$$= \frac{\delta_{\{m\},\{m'\}}}{d^2_{\{m\}}}\left|\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\,\mathrm{tr}(\pi_{\{m\}}(1_N)g_{\{m\}}(\beta_{\{m\}}))\right|^2. \qquad (VI.10)$$

Here, in the third step we have used the fact that $\beta_{\{m\}}, \Theta\beta_{\{m'\}}$ are supported in disjoint domains of space–time, the time reflection invariance of the measure and its maximal clustering property of the measure (non-overlapping loops are non-interacting). In the last step we used the fact that the integral over $g_{\{m\}}(\beta_{\{m\}})_{AB}$ results in a constant matrix, M_{AB} say, which, by inspection of (IV.12) is a linear combination of projectors onto representation spaces of irreducible representations, partially contracted as to match the index structure of $\pi_{\{m\}}$. So M is a linear combination of matrices of the form $\sigma'_{AB} = \sigma_{C,A;C,B}(1_N)$, where σ is an irreducible projector. Now using the fact that $\sigma'\pi_{\{m\}}(1_N) = \pi_{\{m\}}(1_N)\sigma'$, that $\pi_{\{m\}}$ is irreducible and that the contraction of tensor products of Kroneckers is again proportional to a tensor product of Kroneckers it follows that $M = \pi_{\{m\}} \times(1_N)\mathrm{tr}(M)/d_{\{m\}}$.

Formula (VI.10) says that

$$\Psi - \sum_{\{m\}} z_{\{m\}}\frac{1}{d_{\{m\}}}\left[\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu\,\mathrm{tr}(\pi_{\{m\}}(1_N)g_{\{m\}}(\beta_{\{m\}}))\right]\chi_{\{m\}}(h_\gamma) \qquad (VI.11)$$

is a null vector. Therefore, our Hilbert space $\mathscr{H}$ is the completion of the linear span of the states $\chi_{\{m\}}(h_\gamma)$ with respect to the Haar measure $d\mu_H$. On the plane, since there is no homotopically non-trivial loop γ, the only state is the constant function $\Psi = 1$ which corresponds precisely to the

trivial quantum theory as obtained via the Hamiltonian formalism. On the cylinder we obtain $\mathscr{H} = L_2(C(G), d\widetilde{\mu}_H)$, where $C(G)$ is the Cartan subgroup of G and $\widetilde{\mu}_H$ is the corresponding effective measure on $C(G)$ induced by the Haar measure μ_H.

Finally, note that, in the final picture, the loop γ probes the connection $\overline{A}$ at time $t = 0$ only. This is is completely analogous to the corresponding construction for the free massless scalar field[6] where the Hilbert space construction can be reduced to the fields at time zero.

Having constructed the Hilbert space, let us now turn to the Hamiltonian. As indicated in section VI A, the Hamiltonian can be obtained as the generator of the Euclidean time translation semi-group. Denote by $\gamma(t) := T(t)\gamma$ the horizontal loop at time t. Now let $\alpha(t) := \gamma(t) \circ \gamma^{-1}$, then we have by the representation property

$$\chi_{\{m\}}(h_{\gamma(t)}) = \mathrm{tr}(\pi_{\{m\}}(h_{\alpha(t)}) \pi_{\{m\}}(h_{\gamma})) \tag{VI.12}$$

so that according to (IV.16) we have that

$$\chi_{\{m\}}(h_{\gamma(t)}) = \left[\int_{\overline{\mathscr{A}/\mathscr{G}}} d\mu \; \chi_{\{m\}}(\alpha(t)) \right] \frac{\chi_{\{m\}}(h_\gamma)}{d_{\{m\}}}. \tag{VI.13}$$

Hence, according to (IV.14)

$$(\chi_{\{m'\}}, T(t)\chi_{\{m\}}) = \exp(-\tfrac{1}{2}\lambda_{\{m\}} g_0^2 L_x t)\, \delta_{\{m'\},\{m\}} \overset{!}{=} (\chi_{\{m'\}}, \exp(-tH)\chi_{\{m\}})$$

and the completeness of the $\chi_{\{m\}}$ allows us to conclude that

$$H = -\frac{g_0^2}{2} L_x \Delta \tag{VI.14}$$

is the configuration representation of the Hamiltonian.

Finally, let us consider the vacuum state. By inspection, it is given by $\Omega = 1$. It is the unique vector annihilated by the Hamiltonian. We therefore expect that the measure is clustering (see Ref. 6, Theorem 19.7.1). Indeed, notice first that finite linear combinations of products of traces of the holonomy around loops form a dense set $\mathscr{D}$ in $L_2(\overline{\mathscr{A}/\mathscr{G}}, d\mu)$ by construction of $\overline{\mathscr{A}/\mathscr{G}}$. Now recall once again that the measure is not interacting in the sense that if Ψ, Ξ are two elements of $\mathscr{D}$ defined through multi-loops lying in disjoint regions of the plane or the cylinder then it follows immediately from (IV.16) that $\langle \Psi \Xi \rangle = \langle \Psi \rangle \langle \Xi \rangle$. Even if Ψ, Ξ are defined through multi-loops which intersect or overlap then there exists a time parameter t_0 such that the multi-loops involved in Ψ and $T(t)\Xi$ lie in disjoint regions of the plane or the cylinder for all $t \geq t_0$. It then follows from the invariance of the measure under time translations that for $t > t_0$ we have

$$\int_0^t ds \langle \Psi T(s)\Xi \rangle = \int_0^{t_0} ds \langle \Psi T(s)\Xi \rangle + \langle \Psi \rangle \langle \Xi \rangle (t - t_0)$$

and since the first term is finite, clustering is immediate.

Thus, as in scalar field theories, Euclidean invariance and reflection positivity have enabled us to construct the Hamiltonian description from the Euclidean. Furthermore, from sections IV and V it follows that for SU(*N*) and U(1) Yang–Mills theories on $\mathbb{R} \times Rl$ and $S^1 \times Rl$, the Hamiltonian theory constructed through this procedure is *exactly* the same as the standard one, constructed ab-initio via canonical quantization.

VII. SUMMARY

The new results of the present paper can be summarized as follows:

(1) We successfully employed the new integration techniques developed in Refs. 11–13 to compute a closed expression for the Wilson loop functionals for Yang–Mills theory in two Euclidean dimensions.

(2) We proposed an extension of the Osterwalder–Schrader framework for gauge theories and showed how to recover the Hilbert space, the Hamiltonian and the vacuum for the Lorentzian theory starting from our Euclidean framework. For two-dimensional Yang–Mills theories on $\mathbb{R} \times \mathbb{R}$ and on $S^1 \times \mathbb{R}$, the resulting quantum theory completely agrees with the one obtained via canonical quantization. Therefore, two-dimensional Yang-Mills theory constitutes another model theory in the framework of constructive quantum field theory.

(3) Our results are manifestly gauge-invariant, geometrically motivated, require only simple mathematical techniques and the resulting quantum theory is manifestly invariant under the classical symmetry generated by area-preserving diffeomorphisms.

How do these results compare with those available in the literature? Let us begin with the Makeenko–Migdal approach. While they formulated differential equations that the Wilson loops have to satisfy, we have derived a general expression for Wilson loops themselves by directly computing the functional integrals. In the intermediate steps we used a lattice regularization. However, in contrast to the more common practice (in lattice gauge theories) of seeking fixed points of the renormalization group, our results for the continuum theories were then obtained by explicitly taking the limits to remove the regulators. Indeed, our general procedure is rather similar to that used in constructive quantum field theory: we began with a fiducial measure μ_0 on our space $\overline{\mathscr{A}/\mathscr{G}}$ of Euclidean paths, introduced an infra-red and an ultraviolet cutoff, evaluated the characteristic functional of the measure and then removed the regulators. Thus, in the end, we were able to show rigorously that the theory exists in the continuum. In particular, our mathematical framework guarantees the existence of the physical measure for the continuum theory (for which the ''fixed point'' arguments of numerical lattice theory do not suffice.)

While the spirit of our approach is the same as that of the mathematical physics literature on the subject, there are some differences as well. Most of these approaches mimic techniques that have been successful in scalar field theories. Thus, generally, one fixes gauge right in the beginning to introduce a vector space structure on $\overline{\mathscr{A}/\mathscr{G}}$ (see, e.g., Ref. 4). Gauge fixing also brings considerable technical simplifications. However, proofs of invariance of the final expressions under gauge transformations and area preserving diffeomorphisms are then often long. Also, in most of this literature, the Wilson loops are computed for non-overlapping loops. Our results are perhaps closest to those of Klimek and Kondracki.[28] Their framework is also manifestly invariant under gauge transformations and area preserving diffeomorphisms. Furthermore, their results (as well as those of the second paper in Ref. 4) imply that their expressions of Wilson loops in the non-overlapping case admit consistent extensions to all loops. However, they restrict themselves to the structure group SU(2) and the relation to lattice gauge theory — and hence to the conventional Yang–Mills theory — is somewhat obscure.

There are several directions in which our results can be extended. We will conclude by mentioning some examples. First, now that closed expressions for Wilson loops are available, it would be very interesting to check if they satisfy the Makeenko–Migdal equations rigorously. Second, our axiomatic framework is incomplete and it would be very desirable to supplement it, e.g., with techniques from Ref. 7. Another direction is suggested by the fact that, for theories discussed here in detail, we expect that the support of the final physical measure is significantly smaller than the full space $\overline{\mathscr{A}/\mathscr{G}}$ with which we began. Rigorous results that provide a good control on the support would be very useful in refining our axiomatic framework. Finally, it would be interesting to extend our Euclidean methods to closed topologies and compare the resulting framework with the gauge fixed framework of Sengupta.[38]

APPENDIX A: YOUNG TABLEAUX

In the main text we encountered the following issue: We had to integrate a tensor product of group factors $\otimes^n g$ with a measure $d\mu = d\mu_H(g)\exp(\beta/N\Re\,\mathrm{tr}(g))$ which is invariant under conjugation. The representation of G corresponding to the n-fold tensor product of the fundamental representation is not irreducible, so let us decompose it into irreducibles

$$\otimes^n g = \oplus_i \pi_i^{(n)}(g)$$

which is possible since G is compact. Now we have that

$$\pi(h)\left[\int_G d\mu(g)\pi(g)\right] = \left[\int_G d\mu(g)\pi(hgh^{-1})\right]\pi(h) = \left[\int_G d\mu(g)\pi(g)\right]\pi(h)$$

so the integral over $\pi(g)$ commutes with the representation (we have used conjugation invariance of the measure in the last step). Accordingly, by Schur's lemma, we conclude that the integral is proportional to the identity since π was supposed to be an irreducible representation. We can compute the constant of proportionality by taking the trace. Therefore we conclude that

$$\int_G d\mu(g)\pi(g) = \frac{\pi(1_N)}{d(\pi)}\int_G d\mu(g)\chi(g), \quad \text{where } \chi(g) = \mathrm{tr}(\pi(g)) \tag{A1}$$

is the character of the representation. This simplifies the group integrals significantly since we only need the character integrals.

Note that what we are doing here is different from what is usually done in the literature:[30,40] Because we want to evaluate the integral non-perturbatively, we cannot use the stronger property of translation invariance of the Haar measure. In case of the Haar measure we simply have[30]

$$\int_G d\mu_H(g)\pi(g) = \delta_{\pi,0}\pi(1_N), \tag{A2}$$

where 0 denotes the trivial representation.

The solution to the problem of how to decompose an arbitrary tensor product of fundamental representations of SU(*N*) into irreducibles can be found, e.g., in Ref. 39 and we just recall the necessary parts of the theory.

Given an n-fold tensor product of the fundamental representation of a group G, consider all possible partitions $\{m\}$ of n into positive integers of decreasing value,

$$n = m_1 + m_2 + \cdots + m_s, \quad \text{where } m_1 \geq m_2 \geq \cdots \geq m_s > 0 .$$

Such a partition defines a so-called *frame* Y (Young diagram) composed of s horizontal rows with m_i boxes in the *i*th row.

Associated with each frame we construct a certain operator acting on the n-fold tensor product representation as follows: Fill the boxes arbitrarily with numbers $B_1, B_2, \ldots, B_n$ where $B_i \in \{1, 2, \ldots, N\}$. Such a filling of the frame is called a *tableau*. Let P denote the subset of the symmetric group of n elements S_n which only permutes the indices *i* of the labels B_i of each row among themselves and similarly Q denotes the subgroup of S_n permuting only the indices in each column among themselves of the given frame. The relevant operator is now given by

$$e_{\{m\},i}^{(n)} := \sum_{q \in Q} \mathrm{sgn}(q)\hat{q} \sum_{p \in P} \hat{p},$$

where *i* labels the filling and sgn(*q*) denotes the sign of the permutation *q*. The action of $\hat{p}$, say, is

$$\hat{p}\, g_{B_1}^{A_1} \cdots g_{B_n}^{A_n} = g_{B_{p(1)}}^{A_1} \cdots g_{B_{p(n)}}^{A_n},$$

that is, it permutes the *indices* of the *subscript* labels B_i. Because of the complete anti-symmetrization in the columns, no diagram has a row longer than *N* boxes, $s \leqslant N$.

It turns out[39] that each of these *symmetrizers* corresponds to an irreducible representation of GL(*N*), U(*N*), and SU(*N*). Symmetrizers corresponding to different frames give rise to inequivalent representations all of those that correspond to different fillings of the same frame are equivalent. However, not all of the symmetrizers for a given frame are linearly independent, a linearly independent set of tableaux, the so-called *standard tableaux* can be constructed as follows: let the indices *i* of a filling always increase in one row from left to right and in each column from top to bottom. The number of these standard tableaux is given by the formula (if $s = 1$, replace the numerator of the fraction by 1)

$$f_{\{m\}}^{(n)} := n! \frac{\Pi_{1 \leqslant i \langle j \leqslant s}(l_i - l_j)}{\Pi_{i=1}^s (l_i!)}, \quad \text{where} \quad l_i := m_i + s - i, \ \ i = 1, \ldots, s \tag{A3}$$

and it is the number of times that the {*m*}th irreducible representation occurs in the decomposition of $\otimes^n g$ into irreducibles. Now let

$$e_{\{m\}}^{(n)} := \sum_{i=1}^{f_{\{m\}}^{(n)}} e_{\{m\},i}^{(n)} \tag{A4}$$

i.e., the sum of the symmetrizers corresponding to the standard tableaux. This object is called the *Young symmetrizer* of the frame {*m*}. One can show that the standard symmetrizers obey the following (quasi) projector property:

$$[e_{\{m\},i}^{(n)} \otimes^n 1_N][e_{\{m'\},j}^{(n)} \otimes^n 1_N] = \delta_{i,j}\delta(\{m\},\{m'\}) \frac{n!}{f_{\{m\}}^{(n)}} e_{\{m\},i}^{(n)},$$

that is, the sum in (A4) is actually direct and

$$p_{\{m\},i}^{(n)} := \frac{f_{\{m\}}^{(n)}}{n!} e_{\{m\},i}^{(n)} \text{ and } p_{\{m\}}^{(n)} := \frac{f_{\{m\}}^{(n)}}{n!} e_{\{m\}}^{(n)}$$

are projectors onto the representation space of the ith of the equivalent irreducible standard representations given by the frame and on their direct sum, respectively.

In particular we have the resolution of the identity

$$\otimes^n 1_N = \oplus_{\{m\}}[p^{(n)}_{\{m\}} \otimes^n 1_N]. \tag{A5}$$

Let us focus on the unitary groups from now on. For the groups SU(N) we have the following formula for the dimension of the $\{m\}$th irreducible representation[40]

$$d_{\{m\}} = \frac{\Pi_{1 \leqslant i < j \leqslant N}(k_i - k_j)}{\Pi_{I=1}^{N-1}(I!)}, \quad \text{where} \quad k_i = m_i + N - i, \quad m_i := 0 \quad \text{for} \quad i > s. \tag{A6}$$

APPENDIX B: U(1) ON THE TORUS

According to the formulas developed in sections III and IV A it is easy to see that the characteristic functional simply becomes

$$\chi(\alpha) = \frac{\int \Pi_\square d\mu_H(g_\square) d\mu_H(g_x) d\mu_H(g_y) \exp\left(-\beta \Sigma_\square [1 - \Re(g_\square)]\right) T_\alpha(g_\square, g_x, g_y) \delta\left(\Pi_\square g_\square, 1\right)}{\int \Pi_\square d\mu_H(g_\square) \exp\left(-\beta \Sigma_\square [1 - \Re(g_\square)]\right) \delta\left(\Pi_\square g_\square, 1\right)}$$

$$= \delta_{l_x,0} \delta_{l_y,0} \lim_{N \to \infty} \frac{\displaystyle\sum_{n=-N}^{N} \left(\frac{I_n}{I_0}\right)^{N_x N_y} \prod_{I=1}^{k} \left(\frac{I_{n+k_I}/I_0}{I_n/I_0}\right)^{|\alpha_I|}}{\displaystyle\sum_{n=-N}^{N} \left(\frac{I_n}{I_0}\right)^{N_x N_y}}, \tag{B1}$$

where we have employed in the second step the Dirichlet formula[41]

$$\delta(g,1) = \sum_{n=-\infty}^{\infty} g^n \tag{B2}$$

and we could interchange the processes of taking the limit and integration since the Wilson action satisfies all the regularity assumptions for the application of that formula. $I_n(\beta) = \int_{-\pi}^{\pi} d\phi/(2\pi) e^{\beta \cos(\phi) + in\phi}$ is the nth modified Bessel function.

Let us write $N_x N_y = \beta g_0^2 V$ and $|\alpha_I| = \beta g_0^2 A(\alpha_I)$ (V is the volume or total area of the torus and $A_I = A(\alpha_I)$ are the areas of the simple non-overlapping homotopically trivial loops of which α is composed) and use the well-known asymptotic properties of the modified Bessel functions[42] in taking the continuum limit $\beta \to \infty$. The result is

$$\chi(\alpha) = \delta_{l_x,0} \delta_{l_y,0} \lim_{N \to \infty} \frac{\displaystyle\sum_{n=-N}^{N} \exp\left(-\frac{g_0^2}{2}\left[n^2 V - \sum_{I=1}^{k} A_I([n+k_I]^2 - n^2)\right]\right)}{\displaystyle\sum_{n=-N}^{N} \exp\left(-\frac{g_0^2}{2} V n^2\right)}$$

$$= \delta_{l_x,0} \delta_{l_y,0} e^{-\frac{g_0^2}{2}\left[\sum_{I=1}^{k} A_I k_I^2 - \frac{1}{V}\left(\sum_{I=1}^{k} k_I A_I\right)^2\right]} \frac{\displaystyle\sum_{n=-\infty}^{\infty} \exp\left(-\frac{V g_0^2}{2}\left(n + \frac{\Sigma_{I=1}^{k} A_I k_I}{V}\right)^2\right)}{\displaystyle\sum_{n=-\infty}^{\infty} \exp\left(-\frac{V g_0^2}{2} n^2\right)}. \tag{B3}$$

Note that the series in numerator and denominator converge absolutely and uniformly to a non-vanishing limit.

Formula (B3) is the exact and complete result. If we could replace the sums by integrals over the real axis then the fraction involved in (C3) would give just the number 1 and we would be left with the exponential factor only. Note that because $V - A_I \geq \Sigma_{J \neq I} A_J$, exponent in the exponential is non-negative:

$$V \sum_I A_I k_I^2 - \left(\sum_I k_I A_I \right)^2 \geq \sum_{I,J \neq I} k_I^2 A_I A_J - 2 \sum_{I < J} k_I k_J A_I A_J = \sum_{I < J} A_I A_J (k_I - k_J)^2 \geq 0 \qquad \text{(B4)}$$

so that this pre-factor alone could possibly be the generating functional of a positive measure (According to the Riesz-Markov theorem one needed to verify that it is a positive linear functional on $\mathscr{H}\mathscr{A}$).

The characteristic functional (B3) has several interesting features, for example:

(1) While the non-interacting measures had exponents that were linear in the areas of the simple loops, for the interacting theory on the torus we obtain a quadratic dependence on the area, thus *violating the area law*! It is an interesting speculation that the interactive nature of the measure is related to the fact that functional integrals with compact time direction are supposed to describe finite temperature field theories. The interaction then comes from the background heat bath and the characteristic functional is the free energy of a canonical ensemble.

(2) The interactive nature of the continuum measure for compact two-dimensional manifolds such as the torus considered here lets us expect that one would observe a phase transition as in Ref. 17 in the limit $N \to \infty$. However, the large N limit of (B3) is beyond the scope of the present paper and we leave a corresponding analysis for future research.

(3) Notice that expression (B4) is invariant under taking complements (that is, $A \to V - A$) if there is only one simple loop, otherwise the simple loop decomposition of the complemented surfaces is different from the original one.

[1] Y. M. Makeenko and A. A. Migdal, Phys. Lett. B **80**, 255 (1979); **97**, 253 (1980); Nucl. Phys. B **153**, 445 (1979); Yad. Fiz. **32**, 838 (1980); **33**, 1639 (1981); Y. M. Makeenko, *ibid.* **33**, 108 (1980); A. A. Migdal, Nucl. Phys. B **265**, 594 (1986).

[2] Y. Nambu, Phys. Lett. B **80**, 372 (1979).

[3] A. M. Polyakov, Phys. Lett. B **82**, 247 (1979).

[4] L. Gross, C. King, and A. Sengupta, Ann. Phys. (N.Y.) **194**, 65 (1989); B. K. Driver, Commun. Math. Phys. **123**, 575 (1989).

[5] N. Bralic, Phys. Rev. D **22**, 3090 (1980).

[6] J. Glimm and A. Jaffe, *Quantum Physics,* 2nd ed. (Springer, New York, 1987).

[7] E. Seiler, ''Gauge theories as a problem of constructive quantum field theory and statistical mechanics,'' *Lecture Notes in Physics* (Springer, Berlin, 1982), Vol. 159, p. 65.

[8] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, and T. Thiemann, ''A manifestly gauge invariant approach to quantum theories of gauge fields,'' in *Geometry of Constrained Dynamical Systems,* edited by J. Charap (Cambridge University Press, Cambridge, 1994).

[9] T. Thiemann, ''An axiomatic approach to constructive quantum gauge field theory,'' in *Symplectic Singularities and Geometry of Gauge Fields* (Banach Center Publications, CITY, 1996), Vol. 39, p. 389.

[10] A. Ashtekar and C. J. Isham, Class. Quantum Grav. **9**, 1069 (1992).

[11] A. Ashtekar and J. Lewandowski, in *Knots and Quantum Gravity,* edited by J. Baez (Oxford University Press, Oxford, 1994).

[12] A. Ashtekar and J. Lewandowski, J. Math. Phys. **36**, 2170 (1995).

[13] J. Baez, Lett. Math. Phys. **31**, 213 (1994); Adv. Math. **117**, 253 (1996).

[14] A. Ashtekar, *Lectures on Non-Perturbative Canonical Gravity,* Notes prepared in collaboration with R.S. Tate (World Scientific, Singapore, 1991).

[15] E. Witten, Commun. Math. Phys. **141**, 153 (1991); J. Geom. Phys. **9**, 303 (1991).

[16] D. Gross and W. Taylor, Nucl. Phys. B **403**, 395 (1993).

[17] M. Douglas and V. A. Kazakov, Phys. Lett. B **319**, 219 (1993).

[18] M. Blau and G. Thomson, ''Lectures on 2d Gauge Theories: Topological Aspects and Path Integral techniques,'' hep-th/9310144.

[19] J. Baez and W. Taylor, Nucl. Phys. B **426**, 53 (1994).
[20] A. Dhar, G. Mandal and S. R. Wadia, Nucl. Phys. B **436**, 487 (1995).
[21] D. Gross and A. Matytsin, Nucl. Phys. B **429**, 50 (1994); **437**, 541 (1995).
[22] S. G. Rajeev, Int. J. Mod. Phys. A **9**, 5583 (1994); Nucl. Phys. B **436**, 487 (1995).
[23] D. Marolf and J. M. Mourão, Commun. Math. Phys. **170**, 583 (1995).
[24] R. Giles, Phys. Rev. D **8**, 2160 (1981).
[25] Y. Yamasaki, *Measures on Infinite Dimensional Spaces* (World Scientific, Singapore, 1985).
[26] A. Ashtekar and J. Lewandowski, J. Geo. Phys. **17**, 191 (1995).
[27] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974).
[28] S. Klimek and W. Kondracki, Commun. Math. Phys. **113**, 389 (1987); V. A. Kazakov, Nucl. Phys. B **283**, (1981).
[29] D. J. Gross and E. Witten, Phys. Rev. D **21**, 446 (1980).
[30] M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, New York, 1983).
[31] E. M. Stein, *Topics in Analysis Related to the Littlewood-Paley Theory* (Princeton University Press, Princeton, 1970).
[32] S. Rajeev, Phys. Lett. B **212**, 203 (1989).
[33] J. E. Hetrick and Yu. Hosotani, Phys. Lett. B **230**, 88 (1989).
[34] S. V. Shabanov, Phys. Lett. B **318**, 323 (1993); Commun. Theor. Phys. **4**, 1 (1995).
[35] T. Thiemann, Acta Cosmologica, **XXI-2**, 145 (1996).
[36] C. Di Bartolo, R. Gambini, and J. Griego, Phys. Rev. D **51**, 502 (1995); J. N. Tavares, Int. J. Mod. Phys. A **9**, 4511 (1994).
[37] M. Hamermesh, *Group Theory and its Applications to Physical Problems* (Addison-Wesley, Reading, 1962).
[38] A. Sengupta, Ann. Phys. (N.Y.) **221**, 17 (1993).
[39] H. Boerner, *Darstellungen von Gruppen* (Springer, Berlin, 1967); H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, 1946).
[40] T. Maekawa, J. Math. Phys. **26**, 1902, 1910 (1985).
[41] R. Courant, *Vorlesungen über Differential und Integralrechnung*, Vols. 1 and 2 (Springer, Berlin, 1971).
[42] S. Flugge, *Mathematische Methoden der Physik*, Vol. 1 (Springer, Berlin, 1979); I. N. Bronstein, K. A. Semendjajew, *Taschenbuch der Mathematik*, 23rd ed. (Harri Deutsch, Thun und Frankfurt/Main, 1987).