Topological Wilson Operators in Path Integrals

Randy S

Abstract This article introduces **topological Wilson operators**, each localized on a 2-dimensional surface in spacetime. These operators can detect some of the 't Hooft operators described in article 93302, in the same sense that the topological 't Hooft operators described in article 82508 can detect some Wilson operators localized on a 1-dimensional manifold. Even though the construction works in discrete spacetime, the smooth-spacetime limit still plays an important role. This article explains its role.

Contents

1 Introduction

Article 22721 gives an overview of Wilson and 't Hooft operators, distinguishing two different types of each based on the nature of the submanifold X of d-dimensional spacetime on which the operator is nominally localized:[1,2]

operator	notation	number of dimensions of X	type of submanifold	this article
Wilson, type 1	W°	1	proper and neat	
Wilson, type 2	$W^\bullet$	2	proper	✓
't Hooft, type 1	T°	$d-3$	proper and neat	
't Hooft, type 2	$T^\bullet$	$d-2$	proper	

Article 22721 reviews the motivation for considering these operators and explains the meaning of the last column.

Other articles in this series describe Wilson operators of type 1 and 't Hooft operators of both types.[3] This article defines Wilson operators of type 2 for any compact connected gauged group G.[4] They have a topological property analogous to the one that 't Hooft operators of type 2 have.[5] They exist only if G has a nontrivial fundamental group $\pi_1(G)$, which is analogous to the fact that 't Hooft operators of type 2 exist only if the gauged group G has a nontrivial center $Z(G)$. Section 23 will construct Wilson operators of type 2, and section 26 will describe the sense in which these topological Wilson operators can detect some 't Hooft operators of type 1.

This article uses the word **operator** for any modification of the integrand of the path integral, like article 09181 does. Article 02242 explains how this relates to the concept of a linear operator on the Hilbert space.

[1] Here, "nominally localized on X" means "localized in an arbitrarily small neighborhood of X."

[2] The names **type 1** and **type 2** and the superscripts $\circ$ and $\bullet$ are not standard.

[3] Article 22721 cites the articles that introduce these operators.

[4] Article 44135 treats the special case $G = U(1)$, using an approach that assumes G is abelian. This article uses a different approach that does not require G to be abelian.

[5] Article 82508 describes the topological property of 't Hooft operators of type 2.

2 Group theory notation and conventions

- If G and H are Lie groups, then $G \simeq H$ means they're **isomorphic** to each other as Lie groups.[6]
- This article uses the multiplicative convention: if g and g' are elements of a group G, then their composition is denoted $gg' \in G$ and called the **product**.[7]
- $Z(G)$ is the **center** of a Lie group G.[6]
- 1_G is the identity element of G.
- For any element $g \in G$, its **conjugacy class** $[g]$ is the set of all elements of the form $h^{-1}gh$ with $h \in G$. These elements are called **conjugations** of g.
- $r(g)$ is the matrix that represents an element $g \in G$ in a representation r.
- $\text{trace}_r(g)$ is the trace of the matrix $r(g)$.
- $\tilde{G}$ is the **universal covering group**[6,8] of a Lie group G.
- $\sigma : \tilde{G} \to G$ is the covering homomorphism, abbreviated **covering map**.[9]
- K is the **kernel** of the covering map. This is the subgroup of $\tilde{G}$ that is mapped to the identity element of G.
- $\tilde{\sigma}$ is a generic lift from G to the covering group $\tilde{G}$. (section 5)
- $\tilde{\sigma}_1$ is the *canonical lift* that will be reviewed in section 5.
- $\tilde{\sigma}_S$ is a surface-dependent lift that will be defined in section 20.
- G_ε is a small neighborhood of the identity element in G (section 15).

[6] Article 92035 explains what this means.

[7] When $G = U(1)$, elements of the universal covering group $\tilde{G}$ are traditionally treated as real numbers composed using addition (instead of multiplication), but that's just a difference in notation and terminology. The definitions and results in this article still apply to that case.

[8] Every connected Lie group has a universal covering group (Knapp (2023), proposition 1.97).

[9] Article 61813 defines *covering map* more generally, and article 92035 defines **homomorphism** for Lie groups.

3 Other notation and conventions

For the rest of this article, G is the gauged group and $\tilde{G}$ is its universal covering group.

- Spacetime is discretized as described in article 46333. ℓ denotes an oriented 1-cell (link), and $\square$ denotes an oriented 2-cell (plaquette).[10]
- $u(\ell)$ is a G-valued link variable.[10]
- $u(x, C)$ is the G-valued holonomy around a closed curve C made of oriented links,[10,11] starting from the basepoint x.
- If X is an oriented manifold, then X^{-1} is the same manifold with the opposite orientation.
- S is a surface, a 2-dimensional submanifold of spacetime. Exception: in section 7, S denotes the action instead.
- $\tilde{\sigma}_S$ is a surface-dependent lift from G to $\tilde{G}$.[12]
- r is a matrix representation of G or $\tilde{G}$, not necessarily faithful.[11]
- T° is an 't Hooft operator.[13] In section 22, T is a maximal tree in the spacetime lattice.
- In section 26, τ is the composition of modifications of the integrand of the path integral.[14] In sections 27-28, τ is a tubular neighborhood.
- $|X|$ is the number of elements in a set X.

---

[10] Section 6
[11] Section 9
[12] Section 14
[13] Section 1
[14] Article 02242

4 Covering groups and representations

Let G be a compact connected Lie group, $\tilde{G}$ its universal covering group, and K the kernel of the covering map. Then K is a discrete subgroup of the center $Z(\tilde{G})$ of $\tilde{G}$.[15] Examples:

- If $G = U(1)$, then $\tilde{G} = \mathbb{R}$ and $K \simeq \mathbb{Z}$.
- If $G = SU(n)/Z(SU(n))$, then $\tilde{G} = SU(n)$ and $K \simeq \mathbb{Z}_n$, the additive group of integers modulo n.

A **unitary representation** r of G is a homomorphism from G to a unitary matrix group M. Every representation of G may be composed with the covering map $\sigma : \tilde{G} \to G$ to get a representation of $\tilde{G}$:

$$\tilde{G} \xrightarrow{\sigma} G \xrightarrow{r} M.$$

If the fundamental group[16] $\pi_1(G)$ is nontrivial, then $\tilde{G}$ also has other unitary representations that cannot be generated in this way from representations of G.[17] These additional representations of $\tilde{G}$ are the representations of interest in this article.

For the rest of this article, G is a compact connected Lie group, and *representation* means *unitary representation.*

[15] Knapp (2023), proposition 1.101

[16] Article 61813 defines **fundamental group**.

[17] A representation of $\tilde{G}$ is sometimes called a **projective representation** of G, but in this article, a *representation* of G is always an ordinary representation by linear transformations of a complex vector space.

5 The concept of a lift

The Wilson operators introduced in this article will be defined by inserting a complex-valued function of the G-valued link variables into the integrand of the path integral, but the function is defined using a representation of $\tilde{G}$. For this to make sense, we need a way to convert a function $\tilde{f}$ of a $\tilde{G}$-valued variable to a function f of a G-valued variable so it makes sense in the integrand of a path integral with gauged group G. We can do that by choosing a map $\tilde{\sigma}: G \to \tilde{G}$ that satisfies

$$\sigma(\tilde{\sigma}(g)) = g \qquad \text{for all } g \in G. \tag{1}$$

Such a map $\tilde{\sigma}$ will be called a **lift** from G to $\tilde{G}$.

If $\tilde{G} \neq G$, then a lift $\tilde{\sigma}$ is not uniquely determined by the condition (1). Article 60496 constructs a lift $\tilde{\sigma}_1$ that satisfies $\tilde{\sigma}_1(1_G) = 1_{\tilde{G}}$ and is continuous in a subset $G_1 \subset G$ with these properties:

- G_1 is a connected neighborhood of 1_G.
- If $g \in G_1$, then $g^{-1} \in G_1$.
- If $g \in G_1$, then $h^{-1}gh \in G_1$ for all $h \in G$.

It also satisfies

$$\tilde{\sigma}_1(h^{-1}gh) = \tilde{\sigma}_1(h^{-1})\tilde{\sigma}_1(g)\tilde{\sigma}_1(h) \qquad \text{for all } g \in G_1 \text{ and } h \in G. \tag{2}$$

The lift $\tilde{\sigma}_1$ will be called the **canonical lift**.

A lift in the narrow sense defined above takes a single element of G as input and returns a single element of the covering group $\tilde{G}$ as output. We can also define a lift that takes a whole curve in G as input and returns a curve in $\tilde{G}$ as output. This is useful because of the **(unique) path lifting property**, which says that if a continuous path in G starts at 1_G, then it has exactly one lift to a continuous path in $\tilde{G}$ that starts at $1_{\tilde{G}}$.[18] The path lifting property is central to the definition of Wilson operators in representations of $\tilde{G}$ that are not representations of G.

[18] Article 60496

6 Link variables and plaquette variables

This article uses the path integral formulation, treating spacetime as discrete so the math is unambiguous.[19] Article 46333 defines the discretization scheme that will be used here. It involves 0-cells (points), which are the endpoints of 1-cells (links), which form the boundaries of 2-cells (plaquettes), which form the boundaries of 3-cells, and so on up to d-cells, whose union is the d-dimensional spacetime manifold. Each k-cell with $k \geq 1$ can have either of two **orientations**.[20] A path made of consistently oriented links will be denoted C, and C^{-1} denotes its **reverse**, defined by reversing both the order of the links in the sequence and the orientation of each individual link. An individual link will be denoted ℓ, and ℓ^{-1} denotes the link with the same endpoints as ℓ but with the opposite orientation. An individual plaquette will be denoted $\square$ even though a plaquette is not necessarily square, and $\square^{-1}$ denotes the plaquette with the same corners as $\square$ but with the opposite orientation.

The **gauge field** is a collection of G-valued **link variables** $u(\ell)$, one for each oriented link ℓ, subject to the constraint $u(\ell^{-1}) = \big(u(\ell)\big)^{-1}$. An assignment of specific values to all the link variables is called a **configuration of the gauge field**. The ordered product of link variables along a path C, starting from a 0-cell x called the **basepoint**,[21] will be denoted $u(x, C)$. An important special case is the ordered product of link variables around the (oriented) boundary of an oriented plaquette $\square$. This is called a **plaquette variable** and will be denoted $u(x, \square)$,[22] where the basepoint x is one of the 0-cells on the boundary of $\square$. The trace of $u(x, \square)$ in a given representation of G will be called a **traced plaquette variable**.

[19] Currently, the only known well-defined nonperturbative constructions of nonabelian Yang-Mills models in 4d spacetime involve discretizing space or spacetime.

[20] Article 46333

[21] If G is nonabelian, then the product is not invariant under permutations of the factors, so we must choose a basepoint to specify where the product starts.

[22] The notation $u(x, \partial\square)$ would be more consistent with the general notation $u(x, C)$, but the risk of misunderstanding seems too low to justify the extra clutter.

7 Review of the path integral

The path integral has the form[23]

$$\Psi'[u]_{t'} \propto \int_{<t'} [du]\; e^{iS[u]}\Psi[u]_t \tag{3}$$

where

- Ψ and Ψ' are the initial and final states,
- $[u]_t$ denotes the set of link variables whose endpoints are both at time t,
- the integral is over of the link variables that have at least one endpoint in the range $\geq t$ and $< t'$ (with no more than one endpoint at time t'),
- each link variable is integrated over the gauged group G.

The action $S[u]$ is a gauge-invariant function of the link variables. Some amount of Wick rotation must be used to make time evolution unitary in the smooth-spacetime limit.[23] With Wick rotation, the magnitude of the complex-valued factor $e^{iS[u]}$ is

$$\left|e^{iS[u]}\right| = e^{-\beta s_E[u]} \tag{4}$$

where $s_E[u]$ is a nonnegative function proportional to the **euclidean action**, and β is a positive real coefficient. Sections 8 and 19 will explain the role of β in this article.

[23] Article 89053

8 The importance of large β

This article uses the path integral formulation, treating spacetime as discrete so the math is unambiguous, but we are ultimately interested in a limit where arbitrary details of that discrete structure don't have any noticeable effects. When G is nonabelian and the number of spacetime dimensions is $d \in \{3, 4\}$, the model used in this article is believed to have a nontrivial smooth-spacetime limit, but taking a strict limit isn't necessary. Just being close to such a limit is usually sufficient for physics because real-world observations have finite resolution anyway. Being close to such a limit requires two things: it requires making the coefficient β in equation (4) large enough so the correlation length is much greater than the discretization scale s, and it requires using only states whose energy is much lower than $1/s$.[24]

When β is large, the path integral is dominated by configurations of the gauge field that are close to minimizing the euclidean action. These will be called **dominant configurations** of the gauge field. The Wilson operators introduced in this article can be defined for any value of β, but the motivation for considering them comes from the properties they have when β is large.[25] Those properties are the focus of this article.

[24] Article 10142

[25] A similar comment applies to 't Hooft operators of type 1 (article 93302).

9 Wilson operators in representations of G

Many operators are defined by inserting a function of the field variables (the integration variables in the path integral) into the integrand of the path integral. One example is the Wilson operator $W_r^\circ(C)$, where C is a **loop** (connected closed curve) in spacetime and r is a representation of G. This operator is defined by a function of the G-valued link variables. This section reviews its definition.

Let C be any loop made of oriented links, and let $|C|$ denote the number of links in C. Let x be an arbitrary **basepoint**, which can be any endpoint of any link in C. Use the abbreviation

$$u(x,C) \equiv u(\ell_1)u(\ell_2)\cdots u\big(\ell_{|C|}\big) \tag{5}$$

where $\ell_1, \ell_2, ..., \ell_{|C|}$ is the sequence of links around C starting at the given basepoint x, and $u(\ell)$ is the link variable associated with the oriented link ℓ. The product (5) is the **holonomy** (an element of G) around the loop C with basepoint x.

Choose any unitary representation r of G, and let $\text{trace}_r(g)$ denote the trace of the matrix $r(g)$ that represents the element $g \in G$. The Wilson operator $W_r^\circ(C)$ is defined by inserting the complex-valued function

$$w_r(C) \equiv \text{trace}_r\big(u(\cdot, C)\big) \tag{6}$$

into the integrand of the path integral. The basepoint doesn't need to be specified because the trace is invariant under cyclic permutations of the factors in (5).[26] This also makes the function (6) invariant under gauge transformations.

We might as well consider only irreducible representations, because the trace over a reducible representation can always be written as a sum over irreducible representations. For the rest of this article, representations are understood to be irreducible and unitary.

[26]The trace is superfluous when $G = U(1)$.

10 Preview of a generalization

The Wilson operators $W_r^\circ(C)$ reviewed in section 9 are defined for any (irreducible) representation r of G. Section 23 will generalize this to representations of the covering group $\tilde{G}$. This section previews that generalization.

If G is not already simply connected, then some representations of $\tilde{G}$ are not representations of G.[27] The path integral is expressed in terms of G-valued link variables, so the generalization must involve a lift from G to $\tilde{G}$, the domain in which the representation r is defined. Equation (6) can be generalized to

$$w_r(C) \equiv \text{trace}_r\Big(\tilde{\sigma}\big(u(\cdot, C)\big)\Big) \tag{7}$$

for some lift $\tilde{\sigma}$, but the value of (7) depends on which lift $\tilde{\sigma}$ we use.[28] The question is how to choose an appropriate lift. Section 12 will compare a few options, and section 23 will use the appropriate option to define what section 1 called a Wilson operator of type 2, denoted $W^\bullet$.

No matter what lift we use, the operator defined by inserting (7) into the integrand of the path integral clearly reduces to $W_r^\circ(C)$ whenever r is also a representation of the original group G. This is clear because if r is a representation of G, then the lift has no effect and so might as well be omitted, leaving (6).

[27] Section 4

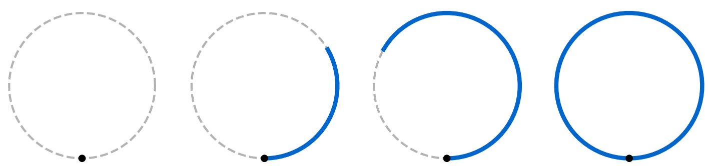
[28] Article 60496

11 Two lifts that don't work

To complete the definition (7), we need to specify the lift $\tilde{\sigma}$. This section highlights two options that don't work as desired.

First consider a single-element lift, one whose input is a single element of G. If $\tilde{G} \neq G$, then any such lift that is defined on all of G must have a discontinuity somewhere on G.[29] The holonomy $u(\cdot, C)$ in equation (7) can be any element of G, so if we tried to define a Wilson operator by applying such a lift to the holonomy as a whole, the resulting Wilson operator would inherit this unnatural discontinuity. Spacetime is treated as discrete, but functions of the G-valued link variables should still be continuous in G. The hamiltonian (energy operator) involves derivatives with respect to those variables,[30] so discontinuities in functions of G could violate the low-energy condition that characterizes the smooth-spacetime limit.[31]

Now consider growing the holonomy $u(x, C)$ as illustrated here:



In these pictures, the dashed circle is the desired loop C, the black dot is the basepoint x, and the blue arcs are successively-larger portions of the loop. In smooth spacetime, the idea is to grow the arc continuously, starting from the basepoint, until it covers the full loop. Parallel transport can be used to assign an element of G to each of those successively-larger arcs. This would define a continuous path in G starting at 1_G, and then the path lifting property would give a continuous path in $\tilde{G}$ starting at $1_{\tilde{G}}$, but this approach has a problem: section 16 will show (using discrete spacetime) that the resulting quantity (7) would depend on the gauge, violating the requirement that $W^\bullet$ should be gauge invariant.

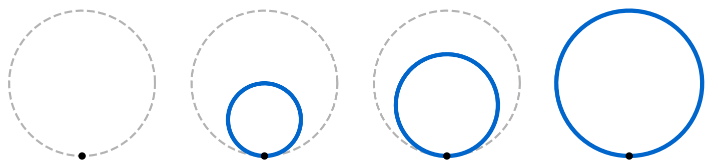
[29]Article 60496

[30]Article 89053

[31]Section 8

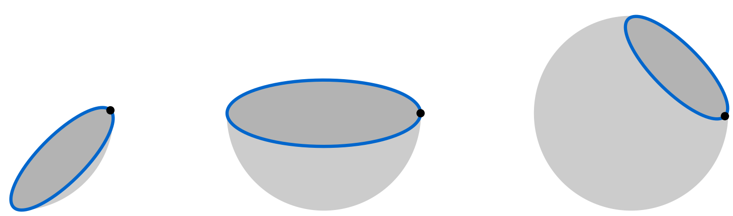
12 How to exploit the path lifting property

The specific approach described at the end of section 11 doesn't work as desired, but the idea of using the path lifting property is on the right track. This sequence of pictures illustrates a better way to grow the holonomy:



In this approach, the curve in spacetime is always closed (always a holonomy). Each holonomy is an element of G, so this continuous sequence of holonomies defines a continuous path in G starting at 1_G. The path lifting property then gives a unique continuous path in the covering group $\tilde{G}$ starting at $1_{\tilde{G}}$. Section 14 will show that the endpoint of the lifted path in $\tilde{G}$ depends only on the surface swept out by the growth of the loop, not on other details of how it grew. Sections 15 and 17-20 will explain how this way of exploiting the path lifting property can be implemented in discrete spacetime. Section 24 will show that it works as desired: the resulting quantity (7) is gauge invariant and independent of the basepoint.

Now consider a loop that grows from a point and then shrinks back to a point, sweeping out a 2-sphere along the way, as illustrated here:



The holonomy around the boundary of each intermediate surface defines a point in G, so this defines a continuous path in G that starts and ends at the same point 1_G. The lifted path in $\tilde{G}$ starts at $1_{\tilde{G}}$ by definition, but it may end at some other element of the kernel K. Sections 26-30 will demonstrate this and explain why it's important.

13 Surfaces that can be swept out by a loop

The approach described in section 12 only works for surfaces that can be swept out using a continuous (or, in discrete spacetime, *practically* continuous) sequence of loops. For surfaces that cannot be grown this way, a corresponding Wilson operator $W^\bullet$ will not be defined.

A disk or 2-sphere embedded in spacetime clearly satisfies the that condition. These are precisely the connected orientable surfaces with genus zero.

More generally, if $\tilde{S}$ is an abstract (not in spacetime) connected orientable surface with genus zero, then the approach in 12 works for any local embedding from $\tilde{S}$ into spacetime,[32] even though the image of a local embedding might not be a surface in the traditional sense. In this generalization, the local embedding (the mapping) is used to pull the gauge field back to the abstract surface, and the growth of the holonomy on that surface defines the path in G. This generalization becomes useful if we want to lift the holonomy around a knotted loop in three-dimensional spacetime, because every knotted loop in a contractible region of three-dimensional smooth spacetime is the boundary of an immersed disk.[33,34,35]

To keep the language concise, most of this article[36] refers only to surfaces in spacetime, specifically to surfaces that are homeomorphic to a disk or to a 2-sphere, even though the constructions and results also apply to any local embedding as described above. Ordinary embedded (not just locally embedded) surfaces in spacetime are the cases of greatest interest.

[32] Article 44113 defines **(topological) embedding**. A **local embedding** is a continuous map that is a (topological) embedding in some neighborhood of each point in the domain. Here, embeddings are allowed to be merely topological (not necessarily smooth) to accommodate sharp bends at the joints between rigid 2-cells.

[33] An **immersion** (defined in Lee (2013), chapter 4) is a smooth local embedding from one smooth manifold to another (Cohen (2023), proposition 3.4).

[34] Hughes and Kim (2020), section 1

[35] This does not contradict the statement that a nontrivial knot has a nonzero *immersed Seifert genus*, because that quantity is defined using only immersions that are embedded along a neighborhood of the boundary (Hughes and Kim (2020), section 1).

[36] Section 31 is an exception.

14 Surface-dependent lift in smooth spacetime

This section shows that the surface-dependent lift defined in section 12 depends only on the surface and does not depend on details of how the loop sweeps out the surface.

Let S be a surface in spacetime that is homeomorphic to a disk, let x be a point on the boundary ∂S, and consider two different ways of sweeping out the surface S by growing a loop C from x to ∂S as illustrated in section 12. This defines two different paths in G that both start at 1_G and end at $u(x, \partial S)$. Let γ_{cat} be the closed path in G obtained by concatenating one of those two paths with the reverse of the other one. The fact that a disk S is contractible implies that every principal G-bundle over S is trivial, so a connection on S may be described using a single local potential A. The holonomy around any loop in S may be expressed using parallel transport with respect to this local potential. The fact that we can morph A continuously to zero (using an overall scale factor) implies that the loop $\gamma_{\text{cat}} \subset G$ is contractible. A contractible loop in G lifts to a loop (closed curve) in the covering group $\tilde{G}$, so both paths from 1_G to $u(x, \partial S)$ in G end at the same point when lifted to $\tilde{G}$. In other words, the lift does not depend on details of how the loop C in spacetime sweeps out the surface.

If S is homeomorphic to a 2-sphere, then it can't necessarily be covered by a single local potential A. To handle this case, form the concatenated path γ_{cat} like before. Omitting an arbitrarily small neighborhood of a point in S would leave a disk, and this omission would change the concatenated path γ_{cat} by only an arbitrarily small amount, so fact that the concatenated path for the disk is contractible implies that the concatenated path for the full sphere is also contractible.

Altogether, this shows that if S is homeomorphic to a disk or to a 2-sphere, then the lift defined in section 12 does not depend on details of how the loop in spacetime sweeps out the surface. The rest of this article uses the notation $\tilde{\sigma}_S\big(u(x, \partial S)\big)$ for that lift. The results derived above justify this notation.

Section 20 will define a discrete-spacetime version of this surface-dependent lift.

15 Discrete version of the path lifting property

The path lifting property says that if a continuous path in G starts at 1_G, then it has exactly one lift to a continuous path in $\tilde{G}$ that starts at $1_{\tilde{G}}$.[37] This section describes a discrete version of the path lifting property: if a *practically* continuous path in G starts at 1_G, then it has a unique natural lift to a *practically* continuous path in $\tilde{G}$ that starts at $1_{\tilde{G}}$.

Define G_ε to be a small neighborhood of the identity element in G that is self-contained under inverses and under conjugation.[38] The neighborhood G_1 mentioned in section 5 also has these properties, but the neighborhood $G_\varepsilon \subset G_1$ is defined to be small enough so that a sequence of products

$$
\begin{aligned}
\gamma_0 &\equiv 1_G \\
\gamma_1 &\equiv g_1 \\
\gamma_2 &\equiv g_1 g_2 \\
\gamma_3 &\equiv g_1 g_2 g_3 \\
&\vdots
\end{aligned}
$$

may be treated as a practically continuous path in G whenever each factor $g_k = \gamma_{k-1}^{-1}\gamma_k$ is in G_ε. Applying $\tilde{\sigma}_1$ to each factor g_k then gives a practically continuous path in $\tilde{G}$, which is the discrete version of the unique path promised by the path lifting property. Explicitly, the sequence of points in $\tilde{G}$ is

$$
\begin{aligned}
\tilde{\gamma}_0 &\equiv \tilde{\sigma}_1(1_G) = 1_{\tilde{G}} \\
\tilde{\gamma}_1 &\equiv \tilde{\sigma}_1(g_1) \\
\tilde{\gamma}_2 &\equiv \tilde{\sigma}_1(g_1)\tilde{\sigma}_1(g_2) \\
\tilde{\gamma}_3 &\equiv \tilde{\sigma}_1(g_1)\tilde{\sigma}_1(g_2)\tilde{\sigma}_1(g_3) \\
&\vdots
\end{aligned}
$$

[37] Section 12

[38] Article 60496 shows that arbitrarily small neighborhoods with these properties exist.

16 Data on the loop itself is not enough

This section shows that if we tried to define the lift using only the gauge field along the loop C (the sequence of pictures in section 11), then the quantity (7) would not be gauge invariant.

Consider a loop C made of a large number of links. We could use a gauge transformation to make the values of all the link variables on the loop C close to 1_G, and in that gauge we could define a lift of $u(\cdot, \partial S)$ by applying $\tilde{\sigma}_1$ to the individual link variables.[39] That's the discrete-spacetime version of what the path lifting property would give in smooth spacetime. This section shows, though, that the resulting lift depends on which gauge we use.

Let $\ell_1, \ell_2, ..., \ell_{|C|}$ be a sequence of oriented links that forms a closed loop C. Use the abbreviation

$$\hat{g} \equiv u(\ell_1)u(\ell_2)\cdots u(\ell_{|C|})$$

for the product of the corresponding link variables $u(\ell_k)$ in the order indicated. The value $\hat{g}$ of the product can be any element of G, not necessarily close to 1_G. Suppose that the number $|C|$ of links in the loop C is large, and let $\hat{\epsilon}$ be an element of G that is close to 1_G and satisfies $\hat{\epsilon}^{|C|} = \hat{g}$. Such an element exists because any element $\hat{g} \in G$ belongs to at least one subgroup of G isomorphic to $U(1)$. Article 00951 shows that we can use a gauge transformation to change the values of those link variables to any list of values whose product is $\hat{g}$.[40] In particular, we can use a gauge transformation to change the values of the link variables to

$$u(\ell_1) = u(\ell_2) = \cdots = u(\ell_{|C|}) = \hat{\epsilon}, \tag{8}$$

so now they're all close to 1_G. Now the sequence

$$1_G,\ u(\ell_1),\ u(\ell_1)u(\ell_2),\ ... \tag{9}$$

[39] This does not rely on β being large.

[40] Article 00951 uses this to make some of them equal to 1_G, but the same method works more generally.

traces out a practically continuous path in G, so the path lifting property[41] suggests defining the lifted value of the product to be

$$\tilde{\sigma}_1\big(u(\ell_1)\big)\tilde{\sigma}_1\big(u(\ell_2)\big)\cdots\tilde{\sigma}_1\big(u(\ell_{|C|})\big) \tag{10}$$

where $\tilde{\sigma}_1$ is the canonical lift defined in section 5. This is the endpoint of the practically continuous lifted path in $\tilde{G}$, starting at the point $1_{\tilde{G}}$.

However, we could also use a gauge transformation to change the values to

$$u(\ell_1) = u(\ell_2) = \cdots = u(\ell_{|C|}) = \epsilon^n\hat{\epsilon} \tag{11}$$

where ϵ is any $|C|$-th root of 1_G in the subgroup generated by $\hat{\epsilon}$, and n is any integer. Now the sequence (9) traces a different path in G: it still starts at 1_G and ends at $\hat{g}$, but it wraps around the $U(1)$ subgroup n times instead of stopping at $\hat{g}$ the first time around. If we use the rule that a (practically) continuous path in G should lift to a (practically) continuous path in $\tilde{G}$, then the lifted value of the product (10) may depend on the integer n.

As an example, let $\tilde{G}$ be the simply-connected group $SU(3)$, let G be the centerless group $SU(3)/\mathbb{Z}_3$, and let $\sigma : \tilde{G} \to G$ be the covering map. Let $\hat{\epsilon} = 1_G$ to make the example as simple as possible, define $\tilde{\epsilon} = \text{diag}(e^{i\phi}, 1, e^{-i\phi}) \in \tilde{G}$ with $\phi = 2\pi/(3|C|)$, and take $\epsilon = \sigma(\tilde{\epsilon})$. Then (11) gives $\hat{g} = 1_G$ for every n, but the value of (10) can be either $\tilde{\epsilon}^0 = 1_{\tilde{G}}$, $\tilde{\epsilon}^{|C|}$, or $\tilde{\epsilon}^{2|C|}$, depending on the value of n.

This shows that if we tried to define the lift using only the gauge field along the loop C (the sequence of pictures in section 11), then a gauge transformation could change the result of the lift even if it doesn't change the value of the product in G.

[41] Article 60496

17 Grafting a plaquette onto a surface

This section shows how to implement the loop-growing approach illustrated by the first sequence of pictures in section 12.

Let S be an immersed oriented surface made of plaquettes in spacetime. Its (oriented) boundary ∂S is a loop made of oriented links. Let x be a basepoint on ∂S, and let $u(x, \partial S)$ be the corresponding holonomy (product of link variables).[42] Let C' be the boundary of a plaquette that shares at least one point with ∂S, let x' be one such point, and let $u(x', C')$ be the corresponding holonomy. Let $\square$ be a plaquette with boundary C', oriented so that any link $\ell' \in C'$ with the same endpoints as a link $\ell \in \partial S$ has the opposite orientation, $\ell' = \ell^{-1}$. Then we can **graft** $\square$ onto S to get a new surface S' and its corresponding holonomy $u(x, \partial S')$. To define this, let C_1 be a part[43] of ∂S that starts at x and ends at x', and let C_2 be the rest of ∂S. Define $u(\cdot, \cdot)$ as in equation (5). Then

$$u(x, C_1)u(x', C')u(x', C_2) = u(x, \partial S'). \tag{12}$$

The example in section 18 should make this evident. Combine (12) with the trivial identity $u(x, C_1)u(x', C_2) = u(x, \partial S)$ to get

$$\boxed{u(x, \partial S') = u(x, \partial S)g} \tag{13}$$

with

$$g \equiv \big(u(x', C_2)\big)^{-1} u(x', C')u(x', C_2).$$

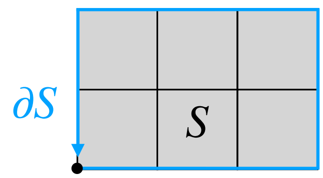
The element $g \in G$ is a conjugation of the plaquette variable $u(x', C')$.

[42] Beware that the notation $u(x, \partial S)$ is ambiguous if the loop ∂S passes through the point x more than once.

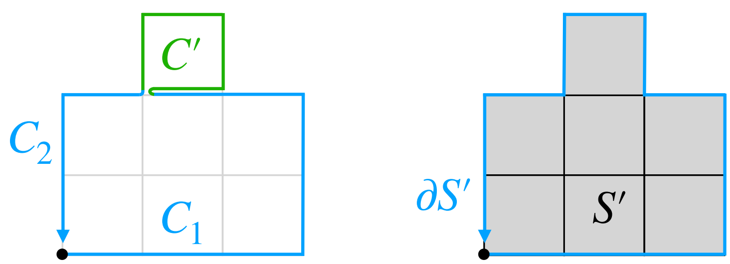
[43] This sentence says *a part* instead of *the part* because the loop ∂S could pass through the point x' more than once.

18 Example

Consider a surface S made from six plaquettes arranged as shown here:[44]



The dot indicates the basepoint x. The blue arrow indicates the orientation of the boundary ∂S, which also defines the orientation of the surface S. Consider a plaquette whose boundary C' shares a link with ∂S but with the opposite orientation, and let x' be one of that link's endpoints. Let C_1 and C_2 be the parts of ∂S before and after the point x'. These pictures illustrate the left and right sides of the identity (12):



Thanks to the identity $u(y,x) = \big(u(x,y)\big)^{-1}$, the last link variable in the product $u(x, C_1)$ cancels the first link variable in the product $u(x', C')$, leaving $u(x, \partial S')$.

[44] This example uses square plaquettes because this makes the pictures easier to draw. More generally, a plaquette may be a polygon whose boundary consists of at least three links (article 46333).

19 Plaquette variables at large β

Define $\tilde{\sigma}_1$ and G_ε as in sections 5 and 15, respectively. When β is sufficiently large,[45] the values of all the plaquette variables are in G_ε for all dominant configurations of the gauge field,[46,47] with one caveat. This section explains the caveat and how it will be managed.

Section 27 will review how an 't Hooft operator $T^\circ(\Gamma)$ is constructed using a prescribed configuration of the gauge field near the codimension 3 submanifold Γ. Suppose that such an operator is present in the integrand of the path integral. Near the smooth-spacetime limit, the path integral is dominated by configurations whose plaquette variables are close to matching a prescribed configuration.[48] If a plaquette $\square$ is close to the core Γ, then the prescribed value – and therefore the plaquette variable $u(\cdot,\square)$ – might not be in G_ε. That's the caveat.

That caveat isn't a problem, because the surface S in $W^\bullet(S)$ should not be that close to Γ anyway. In smooth spacetime, observables can't really be localized on lower-dimensional submanifolds of spacetime like S or Γ. They can only be localized in an arbitrarily small neighborhood of such a submanifold. When a model is defined by treating spacetime as discrete, that arbitrarily small neighborhood becomes a finite keep-out zone: if $W^\bullet(S)$ and $T^\circ(\Gamma)$ are both present in the integrand of a path integral, then S should be far enough from Γ.

In the definition of $T^\circ(\Gamma)$, we can take the keep-out zone to be thick enough (in the directions transverse to Γ) so that the prescribed values are in G_ε for any $\square$ near the boundary of the keep-out zone. Then, when β is sufficiently large, the plaquette variables $u(\cdot,\square)$ will also be in G_ε for all dominant configurations of the gauge field whenever $\square$ is outside the keep-out zone. The surface S in $W^\bullet(S)$ is outside the keep-out zone, so that property holds for $\square\in S$. That's important because $\tilde{\sigma}_1\big(u(\cdot,\square)\big)$ is a continuous function of $u(\cdot,\square)$ as long as $u(\cdot,\square)$ is in G_ε.

[45] Section 8

[46] Article 89053

[47] Section 8 defined *dominant configurations*.

[48] Section 27

20 Surface-dependent lift in discrete spacetime

This section extends the definition of the surface-dependent lift $\tilde{\sigma}_S$ in section 14 from smooth spacetime to discrete spacetime. The definition works as desired as long as each plaquette variable on S is in the neighborhood G_ε defined in section 15. When β is large enough, the path integral is dominated by configurations of the gauge field for which this condition is satisfied.[49]

Let S be an oriented surface homeomorphic to a disk or to a 2-sphere. Let $\square_1, \square_2, ..., \square_{|S|}$ be the list of oriented plaquettes in S, ordered so that the first k plaquettes in the sequence form a surface S_k homeomorphic to a disk for each $1 \leq k < |S|$. Equation (13) shows that if the plaquette variables are all in G_ε, then the holonomy around the boundary changes only slightly from one increment to the next. In other words, the sequence of holonomies

$$1_G,\ u(x, \partial S_1),\ u(x, \partial S_2),\ u(x, \partial S_3),\ ...$$

forms a practically continuous path in G, as illustrated in section 12. According to equation (13), this path ends at the point[50]

$$u(x, \partial S) = g_1 g_2 \cdots g_{|S|} \tag{14}$$

where g_k is an appropriately conjugated plaquette variable for $\square_k$. Now we can invoke a discrete version of the path lifting property:[51] The lift $\tilde{\sigma}_S$ is defined to be the endpoint of that path in $\tilde{G}$. Explicitly:

$$\boxed{\tilde{\sigma}_S\big(u(x, \partial S)\big) \equiv \tilde{\sigma}_1(g_1)\tilde{\sigma}_1(g_2) \cdots \tilde{\sigma}_1\big(g_{|S|}\big)} \tag{15}$$

where $\tilde{\sigma}_1$ is the canonical lift that was defined in section 5. Section 21 will show that the right side of (15) is independent of how the plaquettes are ordered as long as every truncation of the sequence leaves a surface homeomorphic to a disk.

[49] Section 19

[50] When G is abelian, equation (14) becomes (the exponential of) a discrete version of **Stokes's theorem**.

[51] Section 12

21 Independence of the plaquette order

This section shows that the lift defined in section 20 does not depend on how the plaquettes on the surface S are ordered, as long as every truncation of the sequence leaves a surface homeomorphic to a disk and all the plaquette variables on S are in G_ε.

First suppose that S is homeomorphic to a disk.[52] Think of the process of assembling the surface one plaquette at a time as the process of following a path in G. Choose a basepoint x, and use the abbreviation

$$\gamma_n \equiv u(x, \partial S_n) \tag{16}$$

for the holonomy around ∂S_n, where S_n is the generalized disk formed by the first n plaquettes. The sequence

$$\gamma_0,\ \gamma_1,\ \gamma_2,\ \ldots,\ \gamma_{|S|} \tag{17}$$

(with $\gamma_0 \equiv 1_G$) is a practically continuous path in G.[53] The goal is to show that the endpoint of the lifted path in $\tilde{G}$, namely $\tilde{\sigma}_S\big(\gamma_{|S|}\big)$, does not depend on the order in which the plaquettes are incorporated.

Section 14 described a smooth-spacetime version of that condition. An approach similar to the one in section 14 can also be used here, but with one new complication: we need to show that all the link variables on S can be morphed continuously to 1_G in such a way that the path γ remains practically continuous throughout the process. In other words, we need the values of all the increments $\gamma_{n-1}^{-1}\gamma_n$ to remain in G_ε, which implies that we need the values of all the plaquette variables on S to remain in G_ε. Morphing each link variable monotonically toward 1_G along the shortest possible path in G (within a $U(1)$ subgroup) does not necessarily satisfy that requirement,[54] but we can satisfy the requirement a different way. Section 22 will show that if S is homeomorphic to a disk, then we can use

[52] The end of this section will address the case where S is homeomorphic to a 2-sphere.

[53] Section 15

[54] For an example, suppose $G = SO(3)$, regarded as the group of rotations in a 3d vector space. Consider a

gauge fixing on a maximal tree to express all the link variables on S in terms of plaquette variables on S. The plaquette variables are initially all in G_ε, so now we can morph them all continuously to 1_G while remaining in G_ε, and by construction this gives a consistent way of morphing the link variables. With that adjustment, the reasoning used in section 14 can also be used here to deduce that $\tilde{\sigma}_S\big(\gamma_{|S|}\big)$ does not depend on the order in which the plaquettes are incorporated when S is homeomorphic to a disk.

Now suppose S is homeomorphic to a 2-sphere. In this case, the plaquette variables on S are not all independent of each other: the value of one of them is determined by the values of the others.[55] Because of that constraint, even if all the plaquette variables are initially in G_ε, we can't necessarily morph them all continuously to 1_G while remaining within G_ε.[56] We can deal with this the same way section section 14 did: removing one plaquette from the 2-sphere leaves a disk, and that changes the concatenated path γ_{cat} only slightly because the plaquette variable is in G_ε. The rest of the argument is the same as in section 14.

Altogether, this shows that if S is homeomorphic to a disk or to a 2-sphere and all the plaquette variables in S are in G_ε, then $\tilde{\sigma}_S\big(\gamma_{|S|}\big)$ does not depend on the order in which the plaquettes are incorporated.

plaquette variable given by a product of four link variables, all equal to the same 90° rotation g. The plaquette variable g^4 is initially 1_G (no rotation), but as the rotation angle morphs monotonically toward zero, the plaquette variable g^4 passes through a 180° rotation, which is not in G_ε.

[55] Section 22

[56] When S is homeomorphic to a 2-sphere, the path γ is closed but not necessarily contractible. If it's not contractible, then we can't morph all the plaquette variables continuously to 1_G while remaining within G_ε.

22 Link variables in terms of plaquette variables

This section derives a result that was used in section 21.

Let U_k be the set of all unoriented k-cells in S: U_0 for the discrete points, U_1 for links, and U_2 for plaquettes. Let $u(U_2)$ be the set of all G-valued plaquette variables associated with the oriented counterparts of plaquettes in U_2, and let $u(U_1)$ be the set of all link variables associated with the directed counterparts of links in U_1.

Let S be any oriented connected surface made of plaquettes that is homeomorphic to a disk or to a 2-sphere. This implies that the links in U_1 form a **planar graph**,[57] so **Euler's formula** holds:[58]

$$|U_2| - |U_1| + |U_0| = \begin{cases} 2 & \text{if } S \simeq \text{2-sphere}, \\ 1 & \text{if } S \simeq \text{disk}, \end{cases}$$

where $|U_k|$ is the number of elements of U_k.

Let T be a **maximal tree** for U_1, defined to be a set of (undirected) links such that any two endpoints of any two links in U_1 are connected to each other by exactly one path made of links in T.

The number of links in a maximal tree is $|T| = |U_0| - 1$, so Euler's formula gives

$$|U_2| - \big(|U_1| - |T|\big) = \begin{cases} 1 & \text{if } S \simeq \text{2-sphere}, \\ 0 & \text{if } S \simeq \text{disk}, \end{cases} \tag{18}$$

This says that the number of links remaining after fixing the values on a maximal tree is no greater than the number of plaquettes, so the values of those link variables are uniquely determined by the G-valued plaquette variables.[59]

[57] https://ncatlab.org/nlab/show/planar+graph

[58] The **Euler characteristic** of a cell complex is the number of even-dimensional cells minus the number of odd-dimensional cells (Hatcher (2001), example 0.2 and text above theorem 2.44). The Euler characteristic depends only on the homotopy type of the cell complex (Hatcher (2001), theorem 2.44). A disk may be represented as a cell complex with a single 2-cell, and a 2-sphere may be represented using one 2-cell and one 0-cell.

[59] Equation (18) also says that if S is homeomorphic to a 2-sphere, then the plaquette variables are not all independent of each other. This is analogous to (a discrete version of) the **Bianchi identity**.

23 Wilson operators in representations of $\tilde{G}$

This section constructs the Wilson operators $W^\bullet$ that were previewed in section 1.

Let r be a representation of the covering group $\tilde{G}$ of the gauged group G. Let S be an oriented surface in spacetime that is homeomorphic either to a disk or to a sphere.[60] Sections 17-18 showed how to express $u(\cdot,\partial S)$ as a product of conjugated plaquette variables. Section 20 defined a lift $\tilde{\sigma}_S$ by applying the canonical lift $\tilde{\sigma}_1$ to each of those factors. The Wilson operator $W_r^\bullet(S)$ is defined by inserting this modification of (6) into the integrand of the path integral:[61]

$$w_r^\bullet(S) \equiv \text{trace}_r\Big(\tilde{\sigma}_S\big(u(\cdot,\partial S)\big)\Big). \tag{19}$$

This was anticipated in section 10, but now the lift $\tilde{\sigma}$ that was left unspecified in that section has finally been specified: it is the surface-dependent lift $\tilde{\sigma}_S$ that was constructed in section 20. A basepoint for ∂S is not specified in (19) because section 24 will show that the quantity (19) is independent of the basepoint.

Section 10 mentioned that $W_r^\bullet(S)$ reduces to $W_r^\circ(\partial S)$ whenever r is a representation of the gauged group G. If r is not a representation of G, then $W_r^\bullet(S)$ may be nontrivial even if the surface S is closed ($\partial S = \varnothing$). Section 26 will explain the significance of this.

[60] Section 13

[61] The idea that Wilson operators for representations of $\tilde{G}$ may be defined at the expense of choosing a surface is mentioned in Okuda (2014), section 2.3; which cites Kapustin and Seiberg (2014), section 2; which cites Witten (1997b), section 9.4; which cites Witten (1997c), section 6.6, which says it's obvious and cites nothing.

24 Gauge invariance and basepoint independence

In hindsight, section 17 shows that each factor g_k in equations (14)-(15) is a conjugated plaquette variable:

$$g_k = h_k^{-1} u(\cdot, \square_k) h_k.$$

The quantity h_k accounts for the trail of links from a point on the kth plaquette to a fixed basepoint $x_0 \in \partial S$. (In section 17, this trail is denoted C_3, and h_k is denoted $u(x, C_3)$.) Each quantity g_k is the product of link variables along a closed path that starts at the basepoint x_0, follows a trail to the plaquette $\square_k$, loops once around that plaquette, and then follows the same trail in reverse to get back to x_0. The fact that the path is closed implies that the effect of any gauge transformation is[62]

$$g_k \to h^{-1} g_k h \tag{20}$$

with the same factor $h \in G$ for all k because these closed paths all start and end at the same point x_0. Equations (2), (15), and (20) show that the resulting effect on $\tilde{\sigma}_S(\cdots)$ is

$$\tilde{\sigma}_S(g_1 g_2 g_3 \cdots) \to \tilde{\sigma}_1(h^{-1}) \tilde{\sigma}_S(g_1 g_2 g_3 \cdots) \tilde{\sigma}_1(h). \tag{21}$$

In the definition of $W^\bullet$, the lift $\tilde{\sigma}_S$ occurs inside a trace, so this shows that $W^\bullet$ is gauge invariant.

Gauge invariance implies invariance under cyclic permutations of the G-valued link variables inside the trace,[63] which in turn implies that (19) is independent of the basepoint.

[62]The effects of the gauge transformation cancel each other at each intermediate point on the path:

$$\cdots u(x_j, x_{j+1}) h(x_{j+1}) h(x_{j+1})^{-1} u(x_{j+1}, x_{j+2}) \cdots = \cdots u(x_j, x_{j+1}) u(x_{j+1}, x_{j+2}) \cdots .$$

In this notation, a link is an ordered pair of points, and the quantity h in (20) is $h(x_0)$.

[63]To deduce this, take the h in (21) to be equal to the first link variable in the product $u(\cdot, \partial S)$. Then $hu(\cdot, \partial S)h^{-1}$ is obtained by moving the first link variable to the end of the product.

25 A topological property

This section shows that the lift defined in section 20 is invariant under a large family of deformations of the surface S, at least when β is large.

Consider two oriented surfaces S and S' in spacetime, both with genus zero and with the same boundary $C \equiv \partial S = \partial S'$. Consider a single 3-cell V in spacetime, and suppose that S and S' are identical almost everywhere except that they each include a different part of the boundary of V. Both surfaces may be grown by incrementally adjoining plaquettes in the same sequence except the plaquettes in V. If the the boundary of V is made of a sufficiently small number of plaquettes and the corresponding plaquette variables are all in G_ε, then the result in section 17 implies that these two sequences define two paths in G that differ from each other only slightly. Article 60496 derives a discrete version of the **homotopy lifting property**, which says that if two paths in G differ from each other only slightly, both starting at 1_G and ending at the same point in G, then the lifted paths must also end at the same point in $\widetilde{G}$:

$$\tilde{\sigma}_{S'}\big(u(x,C)\big) = \tilde{\sigma}_S\big(u(x,C)\big). \tag{22}$$

This shows that $W^\bullet(S)$ is invariant under any series of small changes S that preserve its boundary and that only involve plaquettes whose associated variables are in G_ε at every increment. The overall change may be arbitrarily large as long as it can be achieved through a series of small changes like the one described above.

26 Detecting 't Hooft operators: meaning

Suppose S is homeomorphic to a 2-sphere. Section 29 will show that if r is a representation of the covering group $\tilde{G}$ but not of G, then $W_r^\bullet(S)$ has a nontrivial effect. This section explains what the effect is.

Recall that this article is using the word *operator* loosely for any modification of the integrand of the path integral.[64] For any two such operators A and B, let $\tau(A,B)$ denote their composition, defined by applying both modifications to the integrand.[65] This is typically not the same as either algebraic product AB or BA. Choose a Wilson operator $W_r^\bullet(S)$ and an 't Hooft operator $T_\rho^\circ(\Gamma)$,[66] where S is any 2-sphere linked with Γ. Choose any initial state at a time that precedes the time interval spanned by S and Γ. Let X denote either of the two operators $W_r^\bullet(S)$ or $\tau\big(W_r^\bullet(S), T_\rho^\circ(\Gamma)\big)$, and let $|X\rangle$ denote the result of using the path integral to evolve the initial state forward to a time after the time interval spanned by S and Γ, assuming that no other operators are present in the integrand. With the normalization convention used in this article,

$$\Big|\tau\big(W_r^\bullet(S), T_\rho^\circ(\Gamma)\big)\Big\rangle = (\text{trace}_r\,\kappa)\,\Big|T_\rho^\circ(\Gamma)\Big\rangle \tag{23}$$

for some κ in the kernel K of the covering map $\tilde{G} \to G$. Section 29 will determine the value of κ. The result depends on the representation r and on the homomorphism $\rho : U(1) \to G$ that characterizes the 't Hooft operator.[67,68]

In the continuous-time limit, equation (23) implies the operator commutation relation $W_r^\bullet(S')T_\rho^\circ(\Gamma) = (\text{trace}_r\,\kappa)T_\rho^\circ(\Gamma)W_r^\bullet(S')$ whenever S' is a disk pierced by Γ and S' and Γ are both confined to one time.[69]

[64] Section 3

[65] Article 02242

[66] Section 1 explained the meaning of the superscripts $\bullet$ and $\circ$.

[67] Article 49708

[68] In this derivation, replacing $T_\rho^\circ(\Gamma)$ with the identity operator gives another way to derive the result $\kappa = 1_{\tilde{G}}$ that was derived in the first half of section 12 using the balloon-deflating method.

[69] Article 09181

27 Describing 't Hooft operators

The 't Hooft operator $T^\circ(\Gamma)$ is defined by modifying the integrand of the path integral in a tubular neighborhood $\tau(\Gamma)$ of Γ. The modification is such that near the smooth-spacetime limit, the path integral is dominated by configurations for which the plaquette variables $u(\cdot,\Box)$ are close to matching a prescribed configuration $\hat{u}(\cdot,\Box)$.[70] For plaquettes $\Box$ in $\tau(\Gamma)$, the prescribed values $\hat{u}(\cdot,\Box)$ are holonomies in a principal $\breve{G}$-bundle in $\tau(\Gamma)$, where $\breve{G}$ is a subgroup of G isomorphic to $U(1)$.[71] For plaquettes not $\tau(\Gamma)$, $\hat{u}(\cdot,\Box)$ is the identity element 1_G of G.

Configurations that dominate the path integral near the smooth-spacetime limit will give the same value of κ in equation (23) for any 2-sphere S (made of plaquettes) linked once with Γ with a given orientation. This S-independence property of κ follows from two things:

- As the smooth-spacetime limit is approached, the approximation $u(\cdot,\Box)\approx\hat{u}(\cdot,\Box)$ becomes progressively better. Near the boundary of $\tau(\Gamma)$, the prescribed values $\hat{u}(\cdot,\Box)$ are close to 1_G, so the values of the plaquette variables associated with closely-spaced plaquettes are nearly equal to each other.

- The prescribed values all belong to this $U(1)$ subgroup of G, so they all commute with each other. When the the plaquette variables all commute with each other (like the prescribed values $\hat{u}(\cdot,\Box)$ do), the product of the plaquette variables over any 2-sphere is the identity element of G. This is a discrete version of the Bianchi identity.

These properties prevent any sudden jumps in the value of κ under slight (few-plaquette) changes of the surface S. The allowed values of κ are discrete, so the lack of sudden jumps implies that κ is invariant under slight changes of S.

Section 29 will use this information to calculate the value of κ.

[70] Article 93302

[71] To avoid a possible ambiguity (article 93302), the tubular neighborhood $\tau(\Gamma)$ to be thick enough to give an accurate value of the holonomy around the perimeter of each plaquette near the boundary of $\tau(\Gamma)$, in the sense that the holonomies are in G_1 (defined in section 5) and would remain in G_1 if the loop were contracted to a point in smooth spacetime. The minimum required thickness depends on the representation (article 36626).

28 Detecting 't Hooft operators: preparation

Section 29 will calculate the value of κ in equation (23). The calculation itself is easy, but justifying the way the calculation is done requires some explanation. The section provides that explanation.

Thanks to the properties of $T^\circ(\Gamma)$ reviewed in section 27, we can calculate the value of κ in equation (23) by supposing the plaquette variables $u(\cdot,\square)$ are equal to the prescribed values $\hat{u}(\cdot,\square)$ and by using a 2-sphere S that is inside $\tau(\Gamma)$ and linked once with Γ. We can choose S to be just inside the boundary of $\tau(\Gamma)$, where the plaquette variable values are in G_ε.[72,73]

The holonomies $\hat{u}(\cdot,\square)$ all commute with each other, so each factor g_k in equations (14)-(15) is invariant under conjugation and is independent of the choice of basepoint. Growing S plaquette-by-plaquette as illustrated in the second half of section 12 corresponds to sequentially adjoining the factors on the right sides of equations (14)-(15). This gives the sequence (17) in G and its lifted version in $\tilde{G}$. The kth point in the the sequence (17) the holonomy around the boundary of S_k, the part of S formed by the first k plaquettes.

The path (17) in G is practically continuous, so we can use a smooth-spacetime method. Instead of a discrete sequence of surfaces S_k, each having one more plaquette than its predecessor, we can think of a continuous sequence of surfaces $S(t)$ parameterized by $0\leq t\leq 1$ with $S(1)$ being the final sphere. As a function of t, the boundary $\partial S(t)$ starts as a point ($t=0$), grows to a loop and sweeps around the sphere until it shrinks back to a point on the opposite side ($t=1$). This defines a continuous path γ in G (the smooth-spacetime version of the sequence (17)), where each point $\gamma(t)$ on the path is the holonomy around $\partial S(t)$. The lifted path in $\tilde{G}$ is uniquely determined by continuity and the initial condition $\tilde{\gamma}(0)=1_{\tilde{G}}$.

As reviewed in section 27, the prescribed values $\hat{u}(\cdot,\square)$ are holonomies defined by a given bundle on the underlying smooth spacetime. This discrete set of holonomies can be reproduced by a single local potential even if the underlying

[72]Section 19

[73]This avoids the possible ambiguity mentioned in footnote 71 in section 27.

bundle is nontrivial,[74] but if we tried to calculate the holonomies around $\partial S(t)$ in smooth spacetime using just that one local potential, then after some value of t, a small change in $S(t)$ would incur a large change in the holonomy – outside the neighborhood $G_1 \subset G$ of 1_G that section 5 used to define the canonical lift. The definition of the Wilson operator $W^\bullet$ uses that canonical lift, so if we want the result obtained by the smooth-spacetime calculation to agree with the one that would be obtained by using the discrete holonomies $\hat{u}(\cdot, \square)$, then (if the bundle is nontrivial) we cannot use just one local potential. We must cover the sphere with at least two overlapping contractible charts and use a different local potential in each one, related to each other by a transition function where the charts overlap. Then the smooth-spacetime calculation can match the discrete calculation while respecting the condition that a small change in $S(t)$ should incur only a small change in the holonomy, where *small* means the change is given by multiplication by an element of G that is close to 1_G.

[74] Article 11617

29 Detecting 't Hooft operators: calculation

This section calculates the value of κ in equation (23) using the approach introduced in section 28.

Cover the 2-sphere S with two contractible charts, say two hemispheres that overlap only on the equator of the sphere. Choose a smooth local potential in each chart so that both local potentials are consistent with a single connection. Let A_1 and A_2 denote these two local potentials. On the equator where A_1 and A_2 are both defined, they must be related by[75,76]

$$A_2 - A_1 = g^{-1}dg \tag{24}$$

where g is a G-valued function on the equator, taking values in the $U(1)$ subgroup of G selected by ρ. This g is the transition function that specifies how trivial bundles over the two charts are glued together to make the G-bundle over S.

The function g in (24) is a map from the equator into G, and its homotopy class characterizes the bundle.[77] In particular, the bundle is trivial if and only if this closed path in G is contractible. The goal is to show that this closed path in G is in the same homotopy class as the closed path γ defined in section 28.

Without changing the bundle, we can continuously morph the local potential in one chart to be zero: $A_1 = 0$.[78] The the local potential in the other chart must be morphed to preserve the condition (24), so if $A_1 = 0$, then

$$A_2 = g^{-1}dg \qquad \text{(on the equator)}. \tag{25}$$

These changes of the local potentials change the path $\gamma(t)$ in G, but they can't change its homotopy class because they are all achieved by continuously morphing the local potentials, with no sudden jumps.

---

[75] Article 76708

[76] This uses the convention in which a local potential is antihermitian (article 49708).

[77] Article 33600

[78] We can do this by continuously reducing an overall constant factor in A_1 from 1 to 0.

Suppose that the holonomies $\gamma(t)$ start in the chart where A_1 is defined.[79] Since $A_1 = 0$, the path $\gamma(t)$ doesn't stray from 1_G until $\partial S(t)$ reaches the equator where A_2 is also defined. Take the function g to be

$$g = e^{i\alpha\mu} \tag{26}$$

where α is the azimuthal angle around the equator, with μ defined as in article 49708. At the value of t that makes $\partial S(t)$ coincide with the equator,[80] the holonomy is still 1_G, because[81]

$$\begin{aligned} \exp\left(\int_{\partial S(t)} A_1\right) &= \exp(0) = 1_G \\ \exp\left(\int_{\partial S(t)} A_2\right) &= \exp\left(2\pi i\,\mu\right) = 1_G. \end{aligned} \tag{27}$$

For all subsequent values of t,[82] the holonomy must be calculated using A_2. The important properties of A_2 are that it has the form (25) on the equator and that must be smooth throughout the chart where it's defined.[83] We can satisfy those conditions by taking

$$A_2 = f(\beta)g^{-1}dg \tag{28}$$

with g given by (26), where $f(\beta)$ is any function of the "latitude" angle β that equals 1 on the equator and goes smoothly to zero at the pole in the middle of the second chart. As $\partial S(t)$ shrinks from the equator to that pole, the value of $\int_{\partial S(t)} A_2$ decreases smoothly to zero, so the path traced out by the shrinking holonomy $\exp\left(\int_{\partial S(t)} A_2\right)$ goes smoothly from 1_G back to 1_G (necessarily departing from 1_G along the way).

[79] Take this chart to be the lower hemisphere in the pictures in section 12.

[80] This corresponds to the second picture in section 12.

[81] Article 49708 explains why μ satisfies (27).

[82] These values of t correspond to the third picture in section 12.

[83] This is the upper hemisphere in the pictures in section 12.

That path $\gamma(t)$ in G, with one point per holonomy (one point per t), is the same as the path traced out by following the transition function around the equator of $S(1)$. To show this, use

$$\begin{aligned}\exp\left(\int_{\partial S(t)} A_2\right) &= \exp\left(\int_{\partial S(t)} f\big(\beta(t)\big)g^{-1}dg\right) && \text{(equation (28))}\\ &= \exp\left(i\mu f\big(\beta(t)\big)\int_0^{2\pi} d\alpha\right) && \text{(equation (26))}\\ &= \exp\left(2\pi i\mu f\big(\beta(t)\big)\right).\end{aligned}$$

As t goes from 0 to 1, $f\big(\beta(t)\big)$) also goes from zero to 1. This shows that the path γ (each point of which is a holonomy) is in the same homotopy class as the transition function (equation (26)) that defines the bundle. The homotopy class is an element of the fundamental group $\pi_1(G)$, and the value of κ in equation (23) is obtained by lifting that path to $\tilde{G}$, so κ is determined by the transition function through the isomorphism between K and the fundamental group $\pi_1(G)$.[84]

In particular, the G-bundle is trivial if and only if that path is contractible in G (if and only if it lifts to a closed loop in $\tilde{G}$),[85] so the κ in equation (23) will be a non-identity element of K if and only if the prescribed configuration[86] that defines the 't Hooft operator $T^\circ(\Gamma)$ is a non-trivial principal G-bundle.[87] These 't Hooft operators are said to have a nonzero **topological charge**.[88,89] They are the ones that can be detected by some of the topological Wilson operators $W^\bullet$.

[84] Møller (2011), lemmas 8.1 and 8.2 and the text between them

[85] Section 30 will describe an example of a noncontractible path in G.

[86] Section 27

[87] Article 49708

[88] Articles 22721 and 49708

[89] Article 22721 mentions another way to parameterize 't Hooft operators, using representations of the Langlands dual group $G^\vee$. According to Kapustin and Witten (2007) (section 6.2, text before equation (6.10)), the topological charge is nonzero if and only if the center of $G^\vee$ is not trivial in the given representation.

30 Example of a noncontractible path

For an example of a noncontractible path in G, take the covering group $\tilde{G}$ to be $SU(n)$, and take $G = SU(n)/K$ where $K \simeq \mathbb{Z}_n$ is the center of $SU(n)$. Let $\tilde{\gamma}$ be the non-closed path in $SU(n)$ defined by

$$\tilde{\gamma}(t) = \text{diag}\big(z(t), z(t), ..., z(t)\big) \qquad\qquad z(t) = e^{2\pi i\, t/n} \qquad\qquad 0 \leq t \leq 1.$$

The covering map sends $\tilde{\gamma}$ to a path γ in G, and this path γ in G is closed because $\tilde{\gamma}(0)$ and $\tilde{\gamma}(1)$ are both in K. If the sequence of holonomies in section 28 forms this closed path γ, then the value of κ in equation (23) is

$$\kappa = \tilde{\gamma}(1) = \text{diag}\big(e^{2\pi i/3},\, e^{2\pi i/3},\, ...,\, e^{2\pi i/3}\big) \in K.$$

31 Nonabelian flux?

Use the abbreviation $\Phi(x,S) \equiv \tilde{\sigma}_S\big(u(x,\partial S)\big)$ for the $\tilde{G}$-valued quantity in (19), and let $[\Phi(x,S)]$ denote its conjugacy class.[90]

When $G = U(1)$, the quantity $\Phi(x,S)$ is gauge invariant and independent of the basepoint. If the covering group $\tilde{G}$ is treated as the additive group $\mathbb{R}$, then $\Phi(x,S)$ has all the right properties to be interpreted as the electromagnetic flux on the surface S.[91]

When G is nonabelian, $\Phi(x,S)$ is neither gauge invariant nor independent of the basepoint x, but its conjugacy class is gauge invariant and independent of the basepoint, so by analogy with the case $G = U(1)$, we might consider calling $[\Phi(x,S)]$ the **flux** on the surface S.[92] In particular, if S is homeomorphic to a 2-sphere, then $\Phi(x,S)$ is an element of the discrete kernel K. This generalizes the fact that when $G = U(1)$, the flux on a closed surface must be an integer multiple of 2π.[93] The analogy is far from perfect, though. The rest of this section describes two ways in which this nonabelian generalization of *flux* deviates from what we might normally expect that word to mean.

Consider a closed orientable surface in spacetime with genus one (a 2-torus). To apply the lift defined in section 12 to such a surface, we need to realize it as a local embedding of a disk.[94] That can be done, but if G is nonabelian, then the lift depends on the embedding (the map) and not just on its image in spacetime. To demonstrate this, realize the torus as the square shown below whose opposite

[90]Section 2

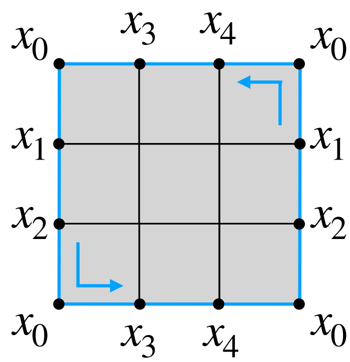
[91]Article 51376 describes another quantity that may be interpreted as the flux on S. In the smooth-spacetime limit, these two quantities should become equal.

[92]When $G = U(1)$, each conjugacy class has only one member (a real number), so we don't need to distinguish between $\Phi(x,S)$ and its conjugacy class in that case.

[93]Sections 25-29 showed that for any G, the flux can be nontrivial (different than the conjugacy class of the identity element) only if the gauge field configuration corresponds to a nontrivial principal G-bundle (and only if $\tilde{G} \neq G$). This phenomenon is familiar when $G = U(1)$ (article 36626).

[94]We could also try to emulate a torus by pinching a 2-sphere so that two opposite disks contact each other face-to-face (an extreme version of a **discocyte** shape, which is the shape of a red blood cell), but this wouldn't accomplish anything that can't already be accomplished with a (distorted) 2-sphere.

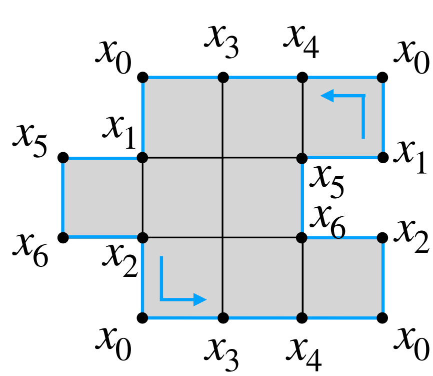
sides are identified with each other in spacetime, as indicated by the labels on the 0-cells around the square's boundary:



Even though the square's image in spacetime is a closed surface (a 2-torus), it still has a nonempty boundary C (the blue outline) as far as the lift defined in section 12 is concerned. Now consider a configuration of the gauge field in which all the link variables are equal to 1_G except $u(x_1, x_2) = g$ and $u(x_3, x_4) = h$ and their oppositely oriented counterparts $u(x_2, x_1) = g^{-1}$ and $u(x_4, x_3) = h^{-1}$. If the boundary is oriented as indicated by the blue arrows, and if we use the square's upper-left corner as the basepoint, then the holonomy is

$$u(x_1, x_2)u(x_3, x_4)u(x_2, x_1)u(x_4, x_3) = ghg^{-1}h^{-1}. \tag{29}$$

The same 2-torus in spacetime can also be realized as the image of a different local embedding, like the one illustrated below:

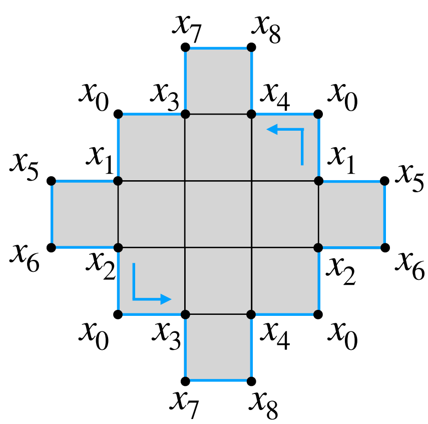


This is almost identical to the previous realization, but the loop C on the surface of the torus in spacetime has been rerouted to avoid both of the oriented links whose endpoints are x_1 and x_2, so now the holonomy is

$$u(x_3, x_4)u(x_4, x_3) = hh^{-1} = 1_G. \tag{30}$$

If G is nonabelian, then the quantities (29) and (30) are typically in different conjugacy classes, so they cannot give the same value for the "flux" on the 2-torus. This demonstrates that the lift depends on the embedding (the map) and not just on its image in spacetime.

Here's another sense in which the nonabelian generalization of *flux* does not behave like we might want. Consider the arrangement shown here:



With the identifications indicated by the labels, this becomes a 2-torus plus two flat wings, each wing consisting of a pair of face-to-face plaquettes with opposite orientations. The corners of one wing are (x_1, x_2, x_6, x_5), and corners of the other wing are (x_3, x_4, x_8, x_7). When $G = U(1)$, the fluxes in each pair of oppositely oriented plaquettes would cancel each other, but they don't cancel when G is nonabelian. This is clear in a configuration where all the link variables are equal to 1_G except $u(x_5, x_6) = g$ and $u(x_7, x_8) = h$ and their oppositely oriented counterparts $u(x_6, x_5) = g^{-1}$ and $u(x_8, x_7) = h^{-1}$. Then the holonomy around the boundary is $ghg^{-1}h^{-1}$, which is typically not equal to 1_G even though the only plaquettes having nonzero individual flux are the ones that occur in oppositely oriented pairs.

32 References

(Open-access items include links.)

Cohen, 2023. "Bundles, Homotopy, and Manifolds" http://virtualmath1.stanford.edu/~ralph/book.pdf

Hatcher, 2001. "Algebraic Topology" https://pi.math.cornell.edu/~hatcher/AT/AT.pdf

Hughes and Kim, 2020. "Immersed Möbius bands in knot complements and representatives of Z_2-homology classes" *Algebr. Geom. Topol.* **20**: 1059-1072, https://arxiv.org/abs/1801.00320

Kapustin and Seiberg, 2014. "Coupling a QFT to a TQFT and duality" *JHEP* **2014(04)**: 001, https://doi.org/10.1007/JHEP04(2014)001

Kapustin and Witten, 2007. "Electric-Magnetic Duality And The Geometric Langlands Program" *Communications in Number Theory and Physics* **1**: 1-236, https://link.intlpress.com/JDetail/1805800346326876162

Knapp, 2023. *Lie Groups Beyond an Introduction.* Anthony W. Knapp, https://www.math.stonybrook.edu/~aknapp/download/Beyond2.pdf

Lee, 2013. *Introduction to Smooth Manifolds (Second Edition).* Springer

Møller, 2011. "The fundamental group and covering spaces" https://arxiv.org/abs/1106.5650

Okuda, 2014. "Line operators in supersymmetric gauge theories and the 2d-4d relation" https://arxiv.org/abs/1412.7126

Witten, 1997b. "Lecture II-9: Wilson loops, 't Hooft loops, and 't Hooft's model of confinement" https://www.math.ias.edu/QFT/spring/witten9.ps

Witten, 1997c. "Lecture II-6: gauge theory in 2 dimensions with self-interacting bosons, teh Wilson line operator, and confinement" https://www.math.ias.edu/QFT/spring/witten6.ps

33 References in this series

Article **00951** (https://cphysics.org/article/00951):
"Complete Gauge Fixing"

Article **02242** (https://cphysics.org/article/02242):
"Localized Operators as States on Boundaries"

Article **09181** (https://cphysics.org/article/09181):
"Topological Operators and Higher-Form Symmetries"

Article **10142** (https://cphysics.org/article/10142):
"Universality and Continuum Limits with Scalar Quantum Fields"

Article **11617** (https://cphysics.org/article/11617):
"Principal Bundles in Lattice Gauge Theory"

Article **22721** (https://cphysics.org/article/22721):
"Overview of Wilson Operators and 't Hooft Operators"

Article **33600** (https://cphysics.org/article/33600):
"Constructing Principal Bundles from Patches"

Article **36626** (https://cphysics.org/article/36626):
"Principal U(1)-Bundles for Codimension 3 't Hooft Operators"

Article **44113** (https://cphysics.org/article/44113):
"Submanifolds and Boundaries"

Article **44135** (https://cphysics.org/article/44135):
"Wilson and 't Hooft Operators for the Quantum Electromagnetic Field at a Single Time"

Article **46333** (https://cphysics.org/article/46333):
"Discretizing Curved Space(time) for Quantum Field Theory"

Article **49708** (https://cphysics.org/article/49708):
"GNO Configurations of Gauge Fields"

Article **51376** (https://cphysics.org/article/51376):
"The Quantum Electromagnetic Field on a Spatial Lattice"

Article **60496** (https://cphysics.org/article/60496):
"Lifting Elements from a Lie Group to its Universal Covering Group"

Article **61813** (https://cphysics.org/article/61813):
"Homotopy, Homotopy Groups, and Covering Spaces"

Article **76708** (https://cphysics.org/article/76708):
"Connections, Local Potentials, and Classical Gauge Fields"

Article **82508** (https://cphysics.org/article/82508):
"Topological 't Hooft Operators in Path Integrals"

Article **89053** (https://cphysics.org/article/89053):
"Quantum Gauge Fields on a Spacetime Lattice"

Article **92035** (https://cphysics.org/article/92035):
"The Topology of Lie Groups: a Collection of Results"

Article **93302** (https://cphysics.org/article/93302):
"'t Hooft Operators from GNO Configurations"