Topological aspect of Chern-Simons quantum field theories

Ninnat Dangniam

Center for Quantum Information and Control,University of New Mexico,Albuquerque, New Mexico, 87131-0001

Contents

Introduction

The Chern-Simons theory is a 2+1-dimensional topological field theory developed and linked to knot invariants (no pun intended) in mathematics by Edward Witten in 1989 [1]. The name comes from the fact that its Lagrangian density is proportional to the Chern-Simons form studied by mathematicians Shiing-Shen Chern and James Harris Simons [2].

In some sense, it is easy to start talking about the Chern-Simons theory by writing the Lagrangian out in full details:

$$\mathcal{L}_{CS} = \frac{k}{4\pi}\epsilon^{jkl}(A_j\partial_k A_l + \frac{2}{3}A_j A_k A_l).$$

But underlying this messy index notation is a beautiful geometric picture of gauge theories, explained in Section I. The Chern-Simons Lagrangian is introduced in Section II. Since the "almost" gauge invariance property and the quantization of the theory involve the path integral formulation of quantum field theory, we only state the results and refer elsewhere for their derivations. Finally, given the connection of the abelian, $U(1)$-Chern-Simons theory to a certain topological invariant, the linking number, we take a look at examples of fractional statistics and topological degeneracy.

I. GEOMETRY OF GAUGE THEORIES

In gauge theories, there are two basic fields that are effected by gauge transformations: the physical fields (i.e. classical or quantum fields) and the gauge fields. Both can be described in the precise language of differential geometry: they are *sections* of particular kinds of *bundles* over the base spacetime *manifold*.

The reader might wonder if such abstraction and generality is really needed, but as we shall see, we are interested in phenomena in topological quantum field theories that arise when the base manifold is topologically nontrivial. In essence, we are doing quantum field theories in a curved spacetime, possibly with boundaries and in dimension other than $3 + 1$.

To develop this language as used by mathematicians would require a book or two. Indeed, the development in this section and the discussion of holonomy in the next section follows the delightful book by John Baez and Javier P.

Muniain, *Gauge Fields, Knots, and Gravity* [3]. But I believe that we already have a good enough geometrical intuition of many notions in differential geometry that would otherwise require and pages of definitions and theorems to justify rigorously. Two of these notions to start with are that of a manifold and a tangent space.

A. Manifolds and tangent spaces

A (smooth) **manifold** M is a space that looks locally Euclidean and on which we can do calculus. A manifold is covered by a collection of Euclidean patches called chart. Formally, a **chart** ϕ_α is a map from an open set U_α to $\mathbb{R}^d$, where d is the dimension of the manifold, roughly the number of parameters required to parametrize it. So we can do a computation in a chart exactly like how we do it in $\mathbb{R}^d$.
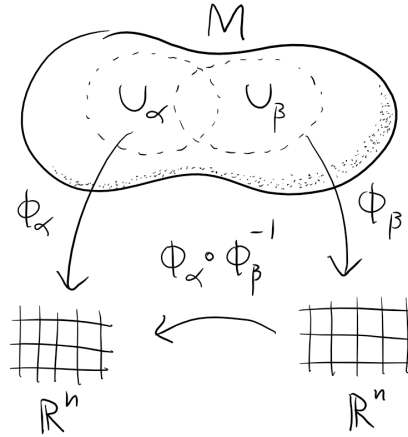

We also require that the transition function between overlapping charts $U_\alpha \cap U_\beta$

$$\phi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1}$$

is infinitely differentiable, that is, a function in C^∞ so that a smooth object in one chart remains smooth in all other overlapping charts. This local chart is required because in general a manifold cannot be covered by just a single chart, just as the globe cannot be represented faithfully by a 2-dimensional map.

The **tangent space** $T_p M$ at point p in an n-dimensional manifold M is an n-dimensional vector space of **tangent vectors**, all possible directions emanated from p. For a flat manifold $M = \mathbb{R}^d$, the manifold and the tangent space are isomorphic and we do not distinguish them. But evoking the sphere as an example again, a tangent vector needs not lie in the manifold in general.
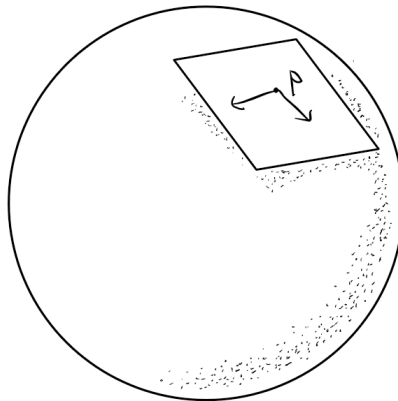
B. Vector bundles

We have said that the sections of particular bundles over a manifold play the role of fields in gauge theories. Consider first a concrete example of a bundle: the **tangent bundle** TM over an n-dimensional manifold M. It is the
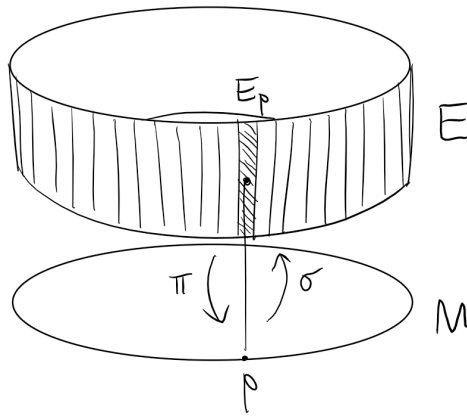
FIG. 1: A cylinder is a trivial fiber bundle with a circle as the base manifold.

collection of tangent spaces at every point $p \in M$.

$$TM = \bigcup_{p \in M} T_p M.$$

Thus a point in TM is a pair (p, v) of a point $p \in M$ and a tangent vector v at that point. So TM is a manifold of dimension $2n$. There is always a surjective projection π from the bundle to M, sending (p, v) to p. So a tangent bundle can be specified as $\pi : TM \to M$.

In a more general bundle

$$\pi : E \to M,$$

the manifold E is called the **total space**, the manifold M is called the **base space**, and π is still a surjective projection from E to M. The pre-image $\pi^{-1}(p)$ is called the **fiber** at p. The fiber E_p now does not have to be a tangent space, but we are still interested in the case when the fiber bundle is a vector space. From this point on, every bundle will be such a bundle, called a **vector bundle**. A right inverse of the projection

$$\sigma : M \to TM$$
$$\pi \circ \sigma = 1,$$

where 1 is the identity map, is called a **section** of the bundle. It assigns to each point in the base space a vector in E_p. For the tangent bundle, a section is nothing more than a **vector field**. But as we alluded to, the physical fields and the gauge fields will be sections of a more general vector bundle.

Example. Gauge fields will turn out to be sections of an **endomorphism bundle** $\text{End}(E)$ of E. It is the bundle with the same base space M as E but with $\text{End}(E)$, the vector space of all linear transformations from E_p to itself, as a fiber at every $p \in M$.

A bundle E is **trivial** if it is just a Cartesian product $E = M \times F$, where F is called a **standard fiber**. Trivial bundles are not so interesting. Nevertheless every bundle that we come across will be **locally trivial**. Local triviality with respect to the standard fiber F means that there exists a neighborhood U of each point $p \in M$ and a bundle isomorphism

$$\phi : E\big|_U \to U \times F,$$

sending each fiber to the trivial bundle $U \times F$. That is to say, E is locally indistinguishable from the trivial bundle.

Going in reverse, a nontrivial bundle can be built up by gluing trivial bundles together. In a G-**bundle** [1], the

———

[1] In some terminology, a G-bundle or a principal G-bundle is not a vector bundle, but its associated vector bundle is. What we describe is the associated vector bundle.
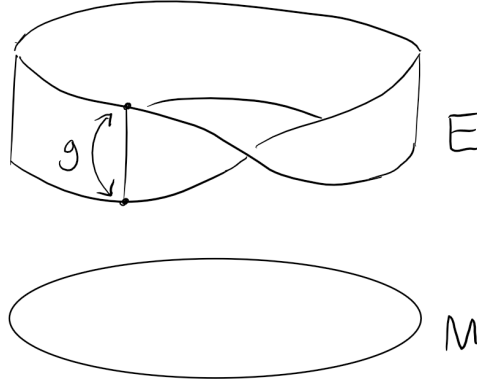
FIG. 2: A möbius strip has the same base manifold as the cylinder but is not a trivial bundle thanks to the twist $g \in \mathbb{Z}_2$ that joins the opposite ends of the fibers.

transition function between overlapping neighborhoods $U_\alpha \cap U_\beta$ in M:

$$\phi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1}$$

where $\phi_\alpha : E\big|_{U_\alpha} \to U_\alpha \times F$ must be an element of the **gauge group** G. (The idea of the transition function between local trivial bundles is analogous to that between charts on a manifold.) The fiber F at each point is now a vector space V on which G acts with a representation ρ. Simultaneous actions of G on every fiber with the same element g are called **global gauge transformations**, whereas simultaneous actions of G with g that depends on each point in the manifold are **local gauge transformations**.

We can finally state the philosophy of gauge theories. In a gauge theory, the physical field is a section of a G-bundle for some group G and if a section s is a solution to the equation of motion (the "physical law"), then gs is also a solution for any local gauge transformation g.

C. Connections and gauge fields

To ensure that the equation of motion is gauge-invariant, it suffices to make the action gauge-invariant. This can be done by "gauging": replacing ordinary derivatives with **connections** (or covariant derivatives). The name comes from the fact that a connection allows us to compare sections at two different points in the manifold and hence allows us to differentiate a section. (Connections also play a central role in the concept of parallel transport, which will be discussed in Section II A.)

A **connection** ∇ on a general vector bundle satisfies the following axioms: Given a function $f \in C^\infty(M)$, tangent vectors (a direction) μ, ν at a point $p \in M$ and a section s, a connection is

1. $C^\infty(M)$-linear in μ

$$\nabla_{f\mu+\nu}s = f\nabla_\mu s + \nabla_\nu s, \tag{1}$$

2. Additive in s

$$\nabla_\mu(s+t) = \nabla_\mu s + \nabla_\mu t, \tag{2}$$

3. And obeys the Leibniz law

$$\nabla_\mu(fs) = (\partial_\mu f)s + f\nabla_\mu s. \tag{3}$$

This is quite abstract, so let us express a connection in coordinates. With the help of a local basis of section $\{e_\mu\}$, every section s can be written as $s = s^\mu e_\mu$. Then the connection of s along the coordinate ν is

$$\begin{aligned}
\nabla_\nu s &= \nabla_\nu(s^\mu e_\mu) \\
&= (\partial_\nu s^\mu)e_\mu + A^\rho{}_{\nu\mu}s^\mu e_\rho \\
&= (\partial_\nu s^\mu + A^\mu{}_{\nu\rho}s^\rho)e_\mu,
\end{aligned}$$

$$\boxed{\nabla_\nu = \partial_\nu s^\mu + A^\mu{}_{\nu\rho}s^\rho} \tag{4}$$

where we write $A^\rho{}_{\nu\mu}e_\rho = \nabla_\nu e_\mu$ and we have only used the Leibniz law. Thus, A specifies how we twist the basis section when we move from one point to another. If the indices μ and ρ are fixed, A is an **endomorphism-valued 1-form** over E or an **End(E)-valued 1-form**. We will say more about forms and operations on them in the next subsection, but for now it means that upon receiving a vector along a coordinate ν, A_ν gives a linear transformation on the fiber E_p. So A itself is also a section, not of E but of End(E).

In a G-bundle, A is called a **gauge field** [2] and it is also effected by gauge transformations. Suppose that a section s is transformed to gs. There is a new connection ∇' such that

$$\nabla'_\mu(gs) = g\nabla_\mu s$$
$$\nabla'_\mu(s) = g\nabla_u(g^{-1}s)$$
$$\partial_\mu s + A'^\mu{}_{\nu\rho}s^\rho e_\mu = g\left(\partial_\mu g^{-1}s + A^\mu{}_{\nu\rho}g^{-1}s^\rho e_\mu\right)$$
$$= \left(gA_\mu g^{-1}\right)(s) + g(\partial_\mu g^{-1})s + \partial_\mu s.$$

Canceling the $\partial_\mu s$ from both sides, we obtain the formula for a gauge transformation

$$\boxed{A'_\mu = gA_\mu g^{-1} + g\partial_\mu g^{-1}}. \tag{5}$$

Observe that a derivative of a Lie group elements is a Lie algebra element. So locally the connection 1-form can also be thought of as a **Lie algebra-valued 1-form** or $\mathfrak{g}$**-valued 1-form** over the G-bundle.

Example. When G is abelian, A is just a number and hence commutes with g, reducing the formula to

$$A'_\mu = A_\mu + g\partial_\mu g^{-1}. \tag{6}$$

QED is an example of abelian gauge theory with $U(1)$ gauge group. Every element of $U(1)$ can be written as $g = e^{-if}$ for some real function f. So the gauge transformation of the potential is

$$A'_\mu = A_\mu + i\partial_\mu f. \tag{7}$$

The new potential differs by a derivative of a function times i. This can be made to agree with what we learned in classical electromagnetism simply by replacing A in the connection with iA.

D. Differential forms

We generalize a bunch of things we know from calculus in 3 dimensions.

The dual space V^* of a vector space V over a field k is the space of all linear functionals that act on a vector to produce a number:

$$\omega : V \to k,$$

a **1-form**. This generalizes the idea of inner product. If V has an inner product "$\cdot$", then there is a vector $w \in V$ such that

$$\omega(v) = w \cdot v.$$

Therefore V and V^* can be identified. (V^* has the same dimension as V.) But there is no canonical choice of an inner product.

───────

[2] For this reason, the gauge field is the same kind of object as the Christoffel symbol in general relativity.

The **wedge product** $\wedge$ on a vector space V generalizes the idea of the cross product. In particular, it is bilinear, associative, and antisymmetric

$$u \wedge v = -v \wedge u, \tag{8}$$

where $u, v \in V$. The antisymmetry means that the wedge product of linearly dependent vectors always vanishes. So we can have at most an n-fold product of vectors in V if V has dimension n. If the vector space of all p-fold product of vectors is denoted by $\Lambda^p V$, the **exterior algebra** ΛV generated by the wedge product on V is graded by p — the degree.

$$\Lambda V = \bigoplus_{p=1}^{n} \Lambda^p V$$

If elements of V are 1-forms, we say that elements of $\Lambda^p V$ are **p-forms**.

An aside: the definition of p-forms by the antisymmetry property still does not quite say what they are explicitly. Another way to think about them is that they are fully antisymmetric tensors on $V^{\otimes p}$ constructed by antisymmetrization $\mathcal{A}$

$$u_1 \wedge u_2 \wedge \cdots \wedge u_n = \mathcal{A}[u_1 \otimes u_2 \otimes \cdots \otimes u_n] = \frac{1}{n!} \sum_{\sigma \in S_n} (-1)^{\text{sgn}\sigma} u_{\sigma(1)} \otimes u_{\sigma(2)} \otimes \cdots \otimes u_{\sigma(n)}, \tag{9}$$

where σ is a permutation in the symmetric group S_n and sgn σ is 1 for an odd permutation and 0 for an even permutation. In the index notation, there is a shorthand using the Levi-Civita symbol

$$\mathcal{A}\left[T_{j_1,j_1,\ldots,j_n}\right] = \frac{1}{n!} \epsilon^{j_1,j_1,\ldots,j_n} T_{j_1,j_1,\ldots,j_n}. \tag{10}$$

To motivate the notion of differential forms, consider the gradient ∇f of a function f in 3 dimensions. (∇ here is not a general connection.) The gradient should be thought of as a 1-form for the following reason: It acts linearly on a vector and produces a number: the directional derivative

$$\nabla f \cdot e_\mu = \partial_\mu f. \tag{11}$$

But more than that, we can have a directional derivative along a vector field s because the gradient is also a $C^\infty(M)$-linear functional. (Recall that a vector field is a section of the tangent bundle, and the space of all sections $\Gamma(E)$ is a $C^\infty(M)$-module, where a module is like a vector space except that the "scalars" can be elements in an arbitrary ring. In this case, it is the ring of C^∞ function on M.) That is, for any smooth function g,

$$\nabla f \cdot (gs) = g(\partial_\mu s). \tag{12}$$

We call such a $C^\infty(M)$-linear functional a **differential 1-form**. All **differential forms** on M can then be generated by the same construction used to generate the exterior algebra.

$$\Omega(M) = \bigoplus_{p=1}^{n} \Omega^p(M)$$

This leads to the idea of the **exterior derivative** d that turns a function f (thought of as a differential 0-form) into a differential 1-form df or more generally, turns a differential p-form into a differential $p+1$ form. The exterior derivative is defined by the following axioms:

1. df is a differential 1-form $df = f_\mu dx^\mu$, where the coordinate 1-forms dx^μ is defined by $dx^\mu(x_\nu) = \delta^\mu{}_\nu$

2. d is scalar-linear.

3. $d(\omega \wedge \mu) = d\omega \wedge \mu - (-1)^p \omega \wedge d\mu$ when $\omega \in \Omega^p(M)$.

4. $d(d\omega) = 0$ for any $\omega \in \Omega(M)$

These are again quite abstract, but they do capture not only the gradient, but also the divergence and the curl and the relation between them in 3 dimensions. These are worked out in Appendix A.
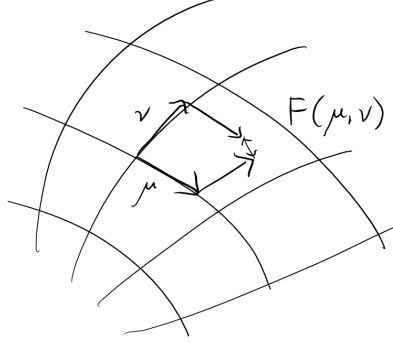
Example. In coordinates, the connection 1-form is

$$A = A^{\mu}{}_{\nu\rho} e_{\mu} \otimes dx^{\rho} \otimes dx^{\nu}, \tag{13}$$

The **curvature** is an antisymmetric 2-form

$$F(\mu, \nu) \equiv \nabla_{\mu} \nabla_{\nu} - \nabla_{\nu} \nabla_{\mu} - \nabla_{[\mu, \nu]}, \tag{14}$$

where $[\mu, \nu]$ in coordinates is the direction of the mixed partial derivatives $\partial_{\mu} \partial_{\nu} - \partial_{\nu} \partial_{\mu}$. It measures the failure of the derivatives ∇_{μ} and ∇_{ν} to commute. The last term corrects for the noncommutativity that arises from the choice of coordinates itself.

To write the curvature in coordinates, let us pick the so-called holonomic basis $\{e_{\mu}\}$ where we do not have the last term. Then

$$
\begin{aligned}
F_{\mu\nu} e_{\rho} &= \nabla_{\mu} \nabla_{\nu} e_{\rho} - \nabla_{\nu} \nabla_{\mu} e_{\rho} \\
&= \nabla_{\mu} (A^{\xi}{}_{\nu\rho} e_{\xi}) - \nabla_{\nu} (A^{\xi}{}_{\mu\rho} e_{\xi}) \\
&= (\partial_{\mu} A^{\xi}{}_{\nu\rho}) e_{\xi} + A^{\eta}{}_{\mu\xi} A^{\xi}{}_{\nu\rho} e_{\eta} - (\partial_{\nu} A^{\xi}{}_{\mu\rho}) e_{\xi} - A^{\eta}{}_{\nu\xi} A^{\xi}{}_{\mu\rho} e_{\eta} \\
&= \left(\partial_{\mu} A^{\xi}{}_{\nu\rho} - \partial_{\nu} A^{\xi}{}_{\mu\rho} + A^{\xi}{}_{\mu\eta} A^{\eta}{}_{\nu\rho} - A^{\xi}{}_{\nu\eta} A^{\eta}{}_{\mu\rho} \right) e_{\xi}
\end{aligned}
$$

Suppressing the first and third indices of A that are waiting for contraction, we get

$$\boxed{F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} + [A_{\mu}, A_{\nu}]}. \tag{15}$$

The commutator does not vanish in general due to the fact that A_{μ} are matrices. It does not vanish, for instance, when the gauge group G is non-abelian. (Recall that A is a Lie algebra-valued 1-form.) The curvature 2-form clearly generalizes the electromagnetic tensor.

We can also see from

$$F = \frac{1}{2} F_{\mu\nu} dx^{\mu} \otimes dx^{\nu}, \tag{16}$$

that the coordinate-free expression of the curvature 2-form is

$$\boxed{F = dA + A \wedge A}. \tag{17}$$

Before we end this section, it will prove necessary to know some algebra of p-forms. The wedge product is antisymmetric for 1-forms (but not $\text{End}(E)$-valued 1-form!) but for p- and q-forms,

$$\omega \wedge \mu = (-1)^{pq} \mu \wedge \omega. \tag{18}$$

An $\text{End}(E)$-valued p-form $S \otimes \omega$ has both the matrix part S and the form part ω. The wedge of product of two $\text{End}(E)$-valued forms can be defined as

$$(S \otimes \omega) \wedge (T \otimes \mu) \equiv ST \otimes (\omega \wedge \nu), \tag{19}$$

where ST is the usual matrix product. The trace of an $\mathrm{End}(E)$-valued form is the trace of the matrix part

$$\mathrm{Tr}(S \otimes \omega) \equiv \mathrm{Tr}(S)\omega. \tag{20}$$

and the exterior derivative operates on the form part

$$d\mathrm{Tr}(S \otimes \omega) = \mathrm{Tr}(S \otimes d\omega). \tag{21}$$

From (18),

$$\begin{aligned}
\mathrm{Tr}(S \otimes \omega \wedge T \otimes \mu) &= (-1)^{pq}\mathrm{Tr}(ST)(\mu \wedge \omega) \\
&= (-1)^{pq}\mathrm{Tr}(TS) \otimes (\mu \wedge \omega) \\
&= (-1)^{pq}\mathrm{Tr}(T \otimes \mu \wedge S \otimes \omega).
\end{aligned} \tag{22}$$

As a result, for any $\mathrm{End}(E)$-value 1-form A,

$$\mathrm{Tr}(A \wedge A) = \frac{1}{2}\mathrm{Tr}([A, A]) = 0 \tag{23}$$

and

$$\mathrm{Tr}(A \wedge A \wedge A \wedge A) = 0. \tag{24}$$

II. CHERN-SIMONS THEORIES

We are all familiar with the source-free electromagnetic Lagrangian density

$$\mathcal{L}_{\mathrm{EM}}(A_\mu) = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu},$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This Lagrangian is not amenable to the coordinate-free formulation with the tools that we have so far since it relies on a metric, and we have not said anything about a metric, for a good reason! Consider instead the somewhat similar-looking but metric-free **2nd Chern form**

$$\mathrm{Tr}(F \wedge F).$$

It turns out that, the action of this Lagrangian is extremized for all choices of A by the Bianchi identity [3] so it gives a rather boring physical theory. But if we instead expand it in terms of the connection 1-form A and use (18) and (24), we find that

$$\begin{aligned}
\mathrm{Tr}(F \wedge F) &= \mathrm{Tr}\left[(dA + A \wedge A) \wedge (dA + A \wedge A)\right] \\
&= \mathrm{Tr}(dA \wedge dA + dA \wedge A \wedge A + A \wedge A \wedge dA + A \wedge A \wedge A \wedge A) \\
&= \mathrm{Tr}(dA \wedge dA + 2A \wedge A \wedge dA) \\
&= d\mathrm{Tr}(A \wedge dA + \frac{2}{3}A \wedge A \wedge A).
\end{aligned}$$

The last line is the exterior derivative of the **Chern-Simons 3-form**

$$\boxed{\mathrm{Tr}(A \wedge dA + \frac{2}{3}A \wedge A \wedge A)} \tag{25}$$

or in coordinates,

$$\epsilon^{jkl}(A_j\partial_k A_l + \frac{2}{3}A_j A_k A_l), \tag{26}$$

where ϵ^{jkl} is the Levi-Civita symbol. In other words, the integral of the 2nd Chern form is the Chern-Simons form.

What happans if we take the Chern-Simons form to be the Lagrangian? First of all, is it gauge invariant? It turns out that the answer is no! [4] Suppose that the action is

$$S_{\mathrm{CS}} = \frac{k}{4\pi} \int_M \mathrm{Tr}(dA \wedge A + A \wedge A \wedge A).$$

Under a gauge transformation $g \in G$, the action gets an additional term

$$S_{CS} \mapsto S_{CS} + 2\pi k n, \tag{27}$$

where n is an integer called the **winding number**. It tells us how many times the gauge transformation goes around the whole group. This fact goes against the philosophy of gauge theories expounded at the end of Section I and seems to ruin any chance to use the Chern-Simons form to construct a physical theory. However, as long as k is also an integer, the amplitude e^{iS} will be gauge invariant for any n. Therefore, we accept the **Chern-Simons theory at integral level** k

$$\boxed{S_{CS} = \frac{k}{4\pi} \int_M \text{Tr}(dA \wedge A + A \wedge A \wedge A)}, \tag{28}$$

to be a physical theory.

A. Wilson loops

What we need now is some gauge-invariant observable for our theory: the Wilson loop.

Another, more geometric meaning of the connection is that it defines a **parallel transport** of a vector u along a path $\gamma(t)$:

$$\nabla_{\gamma'(t)} u = 0 \tag{29}$$

for every t. The picture is that we are walking along the manifold while holding a vector and trying our best to keep the vector pointing in the same direction, but the effort could be spoiled by the curvature of the connection. Take a walk on a sphere as an example. If we walk from the north pole to the south pole along a great circle, then return on a different great circle, the vector ends up rotating even though we are back to the starting point.
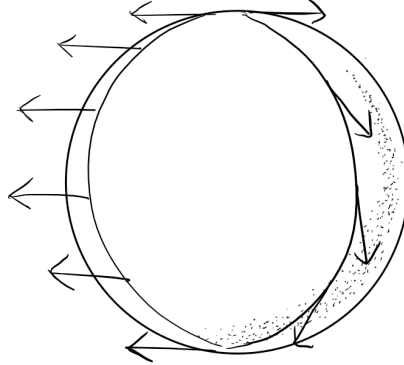
How much the vector is rotated is measured by the holonomy. The **holonomy** $H(\gamma, A)$ on a G-bundle along a path $\gamma(t)$ is the group element in G that maps the vector at the initial time to the vector at the final time. Writing out (29)

$$\partial_{\gamma'} u^\mu + A_{\gamma'} u = 0$$
$$\gamma'(t) \left(\partial_{\gamma'} u^\mu + A_{\gamma'} u \right) = 0$$
$$\frac{du(t)}{dt} + \gamma'(t) A_{\gamma'} u = 0. \tag{30}$$

Since $A(t) \equiv \gamma'(t) A_{\gamma'}$ at different time may not commute, the solution to this is

$$u(t) = P \exp\left[-\int_0^t d\tau A(\tau) \right] u(0), \tag{31}$$

where P signifies the path-ordered exponential

$$P \exp\left[-\int_0^t A(\tau) d\tau \right] = \sum_{n=0}^\infty \frac{(-1)^n}{n!} P \left[\int_0^t d\tau A(\tau) \right]^n, \tag{32}$$

or more compactly,

$$u(t) = Pe^{\int d\gamma \cdot A(t)} u(0). \tag{33}$$

So $Pe^{\int d\gamma \cdot A(t)}$ is the holonomy. (Recall that A takes value in the Lie algebra of G, so the holonomy is indeed the group element.)

Suppose that $u(t)$ is the solution to (30) above. Under a local gauge transformation $g(\gamma(t))$,

$$\begin{aligned}
\frac{d\tilde{u}}{dt} \equiv \frac{d}{dt}(gu) &= \frac{dg(\gamma(t))}{dt}u + g\frac{du(t)}{dt} \\
&= \gamma'(t)(\partial_{\gamma'}g)u - \gamma'(t)gA_{\gamma'}u,
\end{aligned}$$

where we have used the original parallel transport equation in the last line.

$$\begin{aligned}
\frac{d\tilde{u}}{dt} &= \gamma'(t)\left((\partial_{\gamma'}g)g^{-1} - gA_{\gamma'}g^{-1}\right)\tilde{u} \\
&= -\left(g(\partial_{\gamma'}g^{-1}) + gA_{\gamma'}g^{-1}\right)\tilde{u} \\
&= -\gamma'(t)\tilde{A}_{\gamma'}u \\
&= -\tilde{A}(t)\tilde{u},
\end{aligned}$$

where $\tilde{A}$ is the gauged transformed connection. The relation

$$H(\gamma, A')g(\gamma(0))u(0) = g(t)u(t) = g(\gamma(t))H(\gamma, A)u(0)$$

implies that the new holonomy is

$$\boxed{H(\gamma, A') = g(\gamma(t))H(\gamma, A)g^{-1}(\gamma(0))}. \tag{34}$$

The **Wilson loop** is defined to be the trace of the holonomy of a loop γ, that is when $\gamma(t) = \gamma(0)$.

$$\boxed{W(\gamma, A) = \text{Tr}[H(\gamma, A)] = \text{Tr}\left[Pe^{\oint d\gamma \cdot A(t)}\right]} \tag{35}$$

Crucially, it is gauge invariant because the trace is invariant under conjugation by the same g.

B. Example: Fractional statistics

So far our discussion has mostly been about the *classical* Chern-Simons theories. As with other gauge theories, the quantization of the Chern-Simons theories is done in the path integral formulation, which is outside the scope of this article. What is more, there is some subtlety in the calculation of the Wilson loop. So we will have to accept a few facts in order to move forward.

For simplicity, we will consider $U(1)$ quantum Chern-Simons theories. In this case, the Chern-Simons form reduces to

$$A \wedge dA = A_j \partial_k A_l dx^j \wedge dx^k \wedge dx^l \tag{36}$$

Since the group is abelian, there is no need for the path ordering in the Wilson loop anymore, and while we are at it, let us change the convention so that the phase has the same form as the Aharonov-Bohm phase in the natural units $\hbar = 1$.

$$W(\gamma, A) = e^{iq \oint_{\gamma'} d\gamma \cdot A}. \tag{37}$$

The vacuum expected value of a Wilson loop can be expressed in the path integral formulation as

$$\langle W(\gamma) \rangle = \frac{\langle 0| W(\gamma) |0\rangle}{\langle 0|0\rangle} = \frac{\int DA\, W(\gamma, A) \exp\left(i \int_M A \wedge dA\right)}{\int DA \exp\left(i \int_M A \wedge dA\right)}, \tag{38}$$
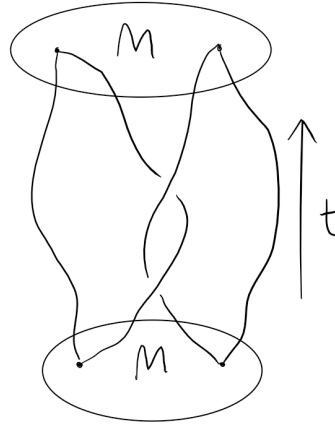
FIG. 3: Winding of the worldlines of two anyon-anti-anyon pairs. The worldlines can also be deformed into the process $T_2^{-1}T_1^{-1}T_2T_1$ showing topological degeneracy in Section II C

where $|0\rangle$ is the vacuum state and DA is the "measure" over the space of gauge fields. The amazing mathematical fact is that this is proportional to a link invariant, the **linking number** L [4], when γ is now allowed to be a collection of loops:

$$\langle W(\gamma)\rangle = \exp\left[\frac{i\pi}{k}\sum_{j,l}q_jq_lL(\gamma_j,\gamma_l)\right],\tag{39}$$

Two loops are linked when one is wrapped around the other. A link in the Chern-Simons theory occurs in 2+1 dimensions. (A curve cannot be wrapped around another without intersecting in 2 dimensions.) That is, it is a linking of worldlines. Suppose that a counterclockwise braiding gives the linking number 1 (and a clockwise braiding gives -1). The vacuum expected value of the Wilson loop in Figure II B will differ from that of a trivial Wilson loop $W(0)$ by

$$\langle W(\gamma)\rangle = \exp\left(\frac{2i\pi q^2}{k}\right)\langle W(0)\rangle\tag{40}$$

This process can be interpreted as the creation of two pairs of **anyons** and their antiparticles from the vacuum followed by a 2π rotation of one particle around the other, and the annihilation into the vacuum again (because we are taking the vacuum expected value).

However, there is clearly still a problem with the formula when $j = l$ because any curve intersects with itself at every point! This problem can be solved by the "framing" prescription. When considering the term $L(\gamma_j,\gamma_j)$, slightly displace γ_j. Call the displaced curve γ_j', and replace $L(\gamma_j,\gamma_j)$ with $L(\gamma_j,\gamma_j')$. This prescription essentially creates a ribbon that can have a self-linking from a twist. (The twist has to be a full 2π rotation. No möbius strip allowed!)
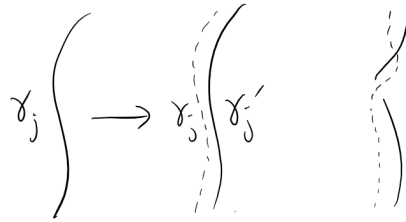

Thus,

$$\langle W(\gamma')\rangle = \exp\left(\frac{i\pi q^2}{k}L(\gamma,\gamma')\right) = e^{\frac{i\pi q^2}{k}} = e^{\frac{i\pi q^2}{k}}\langle W(0)\rangle.\tag{41}$$

The additional phase from such a twist is $\theta = \pi q^2/k$. When interpreted as an anyon rotating about itself, this gives the anyon a spin $q^2/(2k)$ because $s = \theta/(2\pi)$. The Chern-Simons theories therefore demonstrate **fractional statistics**.

C. Example: Topological degeneracy

Suppose now that the spatial manifold is a torus. Processes that start in a vacuum state and end in a vacuum state commute with the Hamiltonian. Ones of these are unitary operators T_1 and T_2 that winds anyons around the torus along the meridian and the longitude respectively.
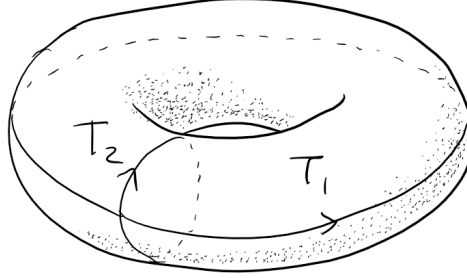
The process $T_1^{-1}T_1$ or $T_2^{-1}T_2$ alone is a loop in spacetime that can be shrunk to a point. But interleaving them creates a link. More precisely, the worldlines of the process that creates an anyon-anti-anyon pair, apply $T_2^{-1}T_1^{-1}T_2T_1$, and annihilate the pair can be deformed into the link in Figure II B. So by (40),

$$T_2^{-1}T_1^{-1}T_2T_1 = e^{2i\theta}$$

or

$$T_2T_1 = e^{2i\theta}T_1T_2. \tag{42}$$

Suppose that $|x\rangle$ is an eigenvector of T_1 with eigenvalue e^{-ix}.

$$T_1(T_2|x\rangle) = e^{-2i\theta}T_2T_1|x\rangle$$
$$= e^{-i(x+2\theta)}(T_2|x\rangle)$$

So $T_2|x\rangle$ is also an eigenvector of T_1 with eigenvalue $e^{-i(x+2\theta)}$, which means that

$$T_2|x\rangle = |x+2\theta\rangle \tag{43}$$

If θ is a rational multiple of π: $\theta = \frac{\pi p}{r}$ with p and r that have no nontrivial common factor, then $\frac{2\pi p}{r}$ has at least r distinct values: $0, 1, 2, \cdots r - 1$. Since T_1 is unitary, all q eigenvectors must be orthogonal. But T_2 is assumed to preserve the ground state, so the ground state must has degeneracy of at least q purely from the topology of the manifold on which the Chern-Simons theory is defined.

This topological degeneracy is the essence of Kitaev's quantum memory [5]. Quantum information encoded in distinct degenerate ground states do not mix up because a random local perturbation is unlikely to send an anyon around the torus completely. Going further, some but not all non-abelian Chern-Simons theories even enable topologically protected universal quantum computation by anyon braiding. This makes sense since braiding of non-abelian anyons gives not just phases but unitary operations. However, not all set of unitary operations are universal. It has been found that anyon braiding in $SU(2)$ Chern-Simons theories at every level k except $k = 1, 2, 4$ allow for universal quantum computation [4]. It is a marvel that one of the most exotic quantum field theories might potentially become a resource for engineering in the future!

[1] Edward Witten, *Quantum field theory and the Jones polynomial*, Communication in Mathematical Physics **121** 351 (1989).

[2] Shiing-Shen Chern and James Harris Simons, *Characteristic forms and geometric invariants*, Annals of Mathematics Second Series **99** 48 (1974).

[3] John Baez and Javier P. Muniain, *Gauge Fields, Knots, and Gravity*, World Scientific (1994).

[4] Jiannis K. Pachos, "Chern-Simons quantum field theories" in *Introduction to Topological Quantum Computation*, Cambridge University Press (2012).

[5] John Preskill, "Lecture Notes for Physics 229: Quantum Information and Computation" www.theory.caltech.edu/preskill/ph229

APPENDIX A: EXTERIOR DERIVATIVES IN 3 DIMENSIONS

The exterior derivative of a 0-form F is the gradient.

$$dF = F_x dx + F_y dy + F_z dz \tag{A1}$$

The exterior derivative of a 1-form $\omega = Adx + Bdy + Cdz$ is

$$
\begin{aligned}
d\omega &= d(Adx + Bdy + Cdz) \\
&= (A_x dx + A_y dy + A_z dz) \wedge dx \\
&\quad + (B_x dx + B_y dy + B_z dz) \wedge dy \\
&\quad + (C_x dx + C_y dy + C_z dz) \wedge dz \\
&= -A_x dx \wedge dy + A_z dz \wedge dx \\
&\quad + B_x dx \wedge dy - B_z dy \wedge dz \\
&\quad - C_x dz \wedge dx + C_y dy \wedge dz \\
&= (C_y - B_z)dy \wedge dz + (A_z - C_x)dz \wedge dx + (B_x - A_y)dx \wedge dy.
\end{aligned}
$$

This returns the definition of the curl if we map the 2-forms to their complementary 1-forms. We can do this because $\Lambda V^1 = V$ and ΛV^2 in 3 dimensions are isomorphic as vector spaces, so we could pretend that every 2-form is a vector! [3]

The exterior derivative of a 2-form $\omega = Ady \wedge dz + Bdz \wedge dx + Cdx \wedge dy$ is

$$
\begin{aligned}
d\omega &= d(Ady \wedge dz + Bdz \wedge dx + Cdx \wedge dy) \\
&= (A_x dx + A_y dy + A_z dz) \wedge dy \wedge dz \\
&\quad + (B_x dx + B_y dy + B_z dz) \wedge dz \wedge dx \\
&\quad + (C_x dx + C_y dy + C_z dz) \wedge dx \wedge dy \\
&= (A_x + B_y + C_z)dx \wedge dy \wedge dz.
\end{aligned}
$$

Again, this is a 3-form. But $\Lambda^3 V$ is 1-dimensional, so this can be identified with the scalar $A_x + B_y + C_z$ and hence the divergence.

The identity $d(d\omega) = 0$ then translates to

$$
\begin{aligned}
\nabla \times (\nabla f) &= 0, \\
\nabla \cdot (\nabla \times A) &= 0.
\end{aligned}
$$

---

[3] The map that realizes this pretension is the Hodge star operator or Hodge dual, which relies on a metric.