22

Topological Field Theory

22.1 What is a Topological Field Theory

We will now consider a special class of gauge theories known as topological field theories. These theories often (but not always) arise as the low energy limit of more complex gauge theories. In general, one expects that at low energies the phase of a gauge theory be either confining or deconfined. While confining phases have (from really good reasons!) attracted much attention, deconfined phases are often regarded as trivial, in the sense that the general expectation is that their vacuum states be unique and the spectrum of low lying states is either massive or massless.

Let us consider a gauge theory whose action on a manifold $\mathcal{M}$ with metric tensor tensor $g_{\mu\nu}(x)$ is

$$S = \int_{\mathcal{M}} d^D x \, \sqrt{g} \, \mathcal{L}(g, A_\mu) \tag{22.1}$$

In section 3.10 we showed that, at the classical level, the the energy-momentum tensor $T^{\mu\nu}(x)$ is the linear response of the action to an infinitesimal change of the local metric,

$$T^{\mu\nu}(x) \equiv \frac{\delta S}{\delta g_{\mu\nu}(x)} \tag{22.2}$$

That a theory is topological means that depends only on the topology of the space in which is defined and, consequently, it is independent of the local properties that depend on the metric, e.g. distances, angles, etc. Therefore, at least at the classical level, the energy-momentum tensor of a topological field theory must vanish identically,

$$T^{\mu\nu} = 0 \tag{22.3}$$

In particular, if the theory is topological, the *energy* (or Hamiltonian) is also zero. Furthermore, if the theory is independent of the metric, it is invariant under *arbitrary* coordinate transformations. Thus, if the theory is a gauge theory, the expectation values of Wilson loops will be independent of the size and shape of the loops. Whether or not a theory of this type can be consistently defined at the quantum level is a subtle problem which we will briefly touch on below.

It turns out that, due to the non-local nature of the observables of a gauge theory, the low energy regime of a theory in its deconfined phase can have non-trivial global properties. In what follows, we will say that a gauge theory is topological if all local excitations are massive (and in fact we will send their mass gaps to infinity). The remaining Hilbert space of states is determined by global properties of the theory, including the topology of the manifold of their space-time. In several cases, the effective action of a topological field theory does not depend on the metric of the space-time, at least at the classical level. In all cases, the observables are non-local objects, Wilson loops and their generalization.

We will focus on two topological field theories: the deconfined phases of discrete gauge theories (particularly the simplest case, $\mathbb{Z}_2$), which exist in any spacetime dimensions $D > 2$, and Chern-Simons gauge theories, which are well understood in 2+1 dimensions.

22.2 Deconfined Phases of Discrete Gauge Theories

Let us consider a simple problem: a $\mathbb{Z}_2$ gauge theory in $D > 2$ space-time dimensions in its deconfined phase. In particular, we will work in the Hamiltonian formulation, c.f. Eq.(18.66), in extreme deconfined limit $\lambda \to \infty$. In this limit, the ground state must satisfy that the each plaquette operator be equal to 1, i.e. no $\mathbb{Z}_2$ flux. These are the *flat* (i.e. no curvature or flux) configurations of the $\mathbb{Z}_2$ gauge field. In this limit, at the local level, this state is satisfied by the configuration that has $\sigma_j^3 = +1$ at every link.

However, this is not is not a gauge invariant invariant state. Furthermore, if the number of sites of bulk of the lattice is N, there are 2^N gauge-equivalent states, obtained by the action of the generator of local gauge transformations $Q(\boldsymbol{r})$ of the $\mathbb{Z}_2$ gauge theory, Eq.(18.53). If the spatial manifold is open, we can fix the gauge locally, e.g. we fix the axial gauge $\sigma_1^3(\boldsymbol{r}) = 1$ on all links along the x_1 axis. Up to a definition of the gauge fields at the boundary of the manifold, this local gauge fixing condition removes all the redundancy.

However, this local gauge fixing condition no longer specifies the state completely if the manifold is closed. For example, if the spatial manifold is a two-torus, the axial local gauge fixing condition does not affects the so-called large gauge transformations, which wrap around a non-contractible loop of the two-torus, such as those shown in Fig.22.2a. In fact, on a closed manifold such as the two-torus, there is a multiplicity of quantum states specified by the solutions of the flat configuration condition. These states are exactly degenerate in the limit in which the system has infinite extent.

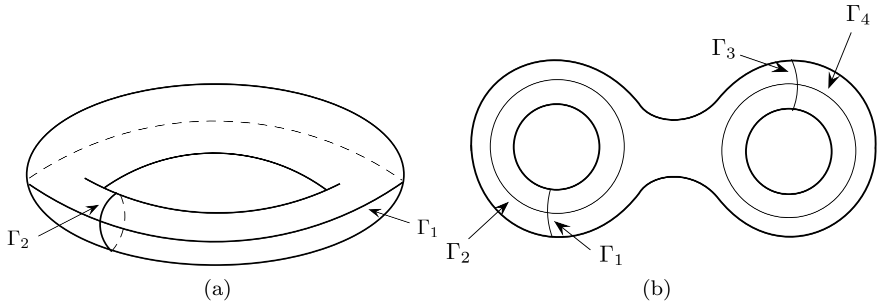

Figure 22.1 a) A two-torus and its two non-contractible loops, Γ_1 and Γ_2; b) The four non-contractible loops of the pretzel.

These degenerate states are specified by the eigenvalues of the Wilson loops along the non-contractible loops of the torus. Indeed, in this ultra-deconfined limit, the eigenvalues of the Wilson loop operators W_{Γ_1} and W_{Γ_2} on the two non-contractible loops Γ_1 and Γ_2 are ± 1. Hence, on a two-torus there are are *four* linearly independent states. On the other hand, a surface of genus 2, a pretzel, has four non-contractible loops (or one-cycles), shown in Fig.22.1b. In general, on a two-dimensional closed surface with of g handles, which in the continuum limit is equivalent to a Riemann surface of genus g, the deconfined phase of the $\mathbb{Z}_2$ gauge theory has 2^{2g} exactly degenerate ground states. In other words, in the low-energy limit of the deconfined phase, the spectrum of the theory is finite-dimensional Hilbert space of dimension 2^{2g}, which grows exponentially with the genus of the surface. This degenerate finite-dimensional Hilbert space has a purely topological origin. The one-cycles shown in Fig.22.1a and b, known as the canonical one-cycles, are linearly independent states of this Hilbert space and constitute a basis of this space.

The magnetic 't Hooft loops and the Wilson loops form a closed algebra. Let $\widetilde{W}[\tilde{\gamma}_i]$, with $i = 1, 2, \ldots, g$, be the 't hooft loops on the canonical one-cycles $\tilde{\gamma}_i$ of a surface of genus g, and let $W[\gamma_j]$, with $j = 1, 2, \ldots, g$, on the canonical one-cycles γ_j. Consider now a Wilson loop operator, which is a product of σ^3 operators on the lop γ_j, and a 't Hooft loop operator, which is a product of σ^1 operators on the links of the lattice pierced by the loop

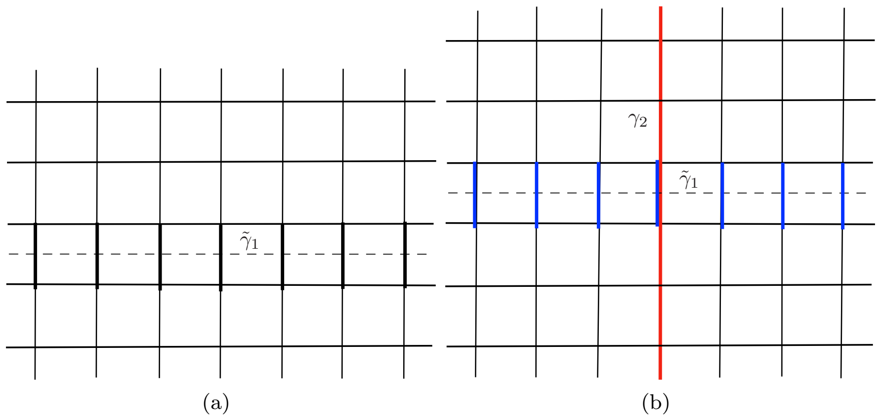

Figure 22.2 a) A magnetic 't Hooft loop $\widetilde{W}[\tilde{\gamma}_1]$ on the non-contractible cycle $\tilde{\gamma}_1$ of the dual lattice; this operator is a product of σ^1 operators on all the links of the lattice pierced by the loop. b) A Wilson loop $W[\gamma_2]$ on the non-contractible loop γ_2 and a 't Hooft loop on the non-contractible loop $\tilde{\gamma}_1$. These operators anticommute with each other.

$\tilde{\gamma}_i$. Since the one-cycles $\tilde{\gamma}_i$ and γ_j cross, they share a link of the lattice as in Fig.22.2b. Hence, these operators anticommute with each other. In general, we have the algebra

$$[W[\gamma_i], W[\gamma_j]] = 0, \quad [\widetilde{W}[\tilde{\gamma}_i], \widetilde{W}[\tilde{\gamma}_i]] = 0, \quad \forall\ i, j \tag{22.4}$$

but

$$\begin{aligned} \{W[\gamma_i], \widetilde{W}[\tilde{\gamma}_j\} &= 0, \quad \text{if } \gamma_i \text{ and } \tilde{\gamma}_j \text{ intersect} \\ [W[\gamma_i], \widetilde{W}[\tilde{\gamma}_j] &= 0 \quad \text{otherwise} \end{aligned} \tag{22.5}$$

Also,

$$W[\gamma_i]^2 = \widetilde{W}[\tilde{\gamma}_i]^2 = I \tag{22.6}$$

where I is the 2×2 identity matrix. It is easy to show that the algebra is the same if the loops are smoothly deformed in the bulk of the system. In other words, the algebra only knows if the loops cross or not. We will see shortly that this is property is related to the concept of braiding.

Since the Wilson loop operators commute with each other, and with the plaquette operator of the Hamiltonian, we can choose the basis of the topological space of states to be the eigenstates of the $2g$ Wilson loops, whose eigenvalues are ± 1. Moreover, since Wilson and 't Hooft loops on cycles belonging to different handles commute with each other, it will suffice to consider the algebra restricted to the two one-cycles of one handle. In that case, there are just two states for each Wilson loop that we will label by $|\pm, i\rangle$, where $i = 1, 2$ labels the two one-cycles of that handle. It is straightforward to see that the algebra of the Wilson and 't Hooft loops implies that

$$\begin{aligned} \widetilde{W}[\tilde{\gamma}_2]|+,1\rangle &= |-,1\rangle, \quad \text{and} \quad \widetilde{W}[\tilde{\gamma}_2]|-,1\rangle = |+,1\rangle \\ \widetilde{W}[\tilde{\gamma}_1]|+,2\rangle &= |-,2\rangle, \quad \text{and} \quad \widetilde{W}[\tilde{\gamma}_1]|-,1\rangle = |+,2\rangle \end{aligned} \tag{22.7}$$

With only minor changes this analysis can be extended to other gauge theories with a discrete abelian gauge group, such as the $\mathbb{Z}_N$ gauge theories. The topological degeneracy is now N^{2g}.

In section 18.8 we discussed that concept of duality. In particular, there we showed that the deconfined phase of the $\mathbb{Z}_2$ gauge theory is the dual of the symmetric phase of the 2+1-dimensional quantum Ising model. The existence of this finite-dimensional topological Hilbert space in the gauge theory seems to contradict the duality mapping. In fact, what we have actually proved is that duality is a relation between local operators of the two theories. The topologically inequivalent subspaces of the gauge theory have the same local content. So, the more precise statement is that duality is blind to the global topology, and that the mapping holds for any topological subspace.

So far we discussed only the case of a theory in 2+1 dimensions. We will now discuss the role of topology in higher dimensions. We will focus on the extreme deconfined limit of the $\mathbb{Z}_2$ gauge theory in 3+1 dimensions with the spatial topology of a three-torus, T^3, and the spacetime manifold is $T^3 \times \mathbb{R}$, where $\mathbb{R}$ is time. A three-torus has three non-contractible one-cycles. Once again the flat configurations can be labeled by the eigenstates of the Wilson loops on the non-contractible one-cycles. Hence, the topological degeneracy is now 8. For a spatial manifold of genus g the degeneracy is 2^{3g}.

Much as in the 2+1-dimensional case, in 3+1 dimensions the 't Hooft operators play an important role. The difference is that while in 2+1 dimensions the 't Hooft loops are strings (essentially, Dirac strings of π flux), in 3+1 dimensions the 't Hooft operators are defined on surfaces, which on T^3 are two-cycles. There are three inequivalent two-cycles that wrap around the three-torus. Now, a Wilson loop on a γ_1 one-cycle will anticommute with the 't Hooft operator of a two-cycle on a surface Σ_{23} pierced by the Wilson loop. Hence, the algebra of Eqs.(22.4)-(22.6) holds here too and the same result follows.

However, there is a caveat in this analysis. While in 2+1 dimensions, a finite but small amount of the kinetic energy term of the gauge theory Hamiltonian, Eq.(18.66), does not affect in an essential way the behavior of Wilson and 't Hooft loops on one-cycles of the the torus, things are different in the deconfined phase in 3+1 dimensions. While the behavior of the Wilson loop is similar, the behavior of the 't Hooft loop is very different. Indeed, even if the kinetic energy term is parametrically very small, a closed 't Hooft loop obeys an area law and, in principle, a 't Hooft operator of a two-cycle will create a state with infinite energy. On the other hand, the area law of the 't Hooft loop turns into a perimeter law once the gauge theory is coupled to a dynamical $\mathbb{Z}_2$ matter field, even one that is very heavy. In that case, the analysis is the same .

We will see in the next section that the algebra of this operators of Eqs. (22.4)-(22.6) is intimately related to the concept of *braiding*, that we will discuss next.

22.3 Chern-Simons Gauge Theories

A particularly important example of a topological field theory is Chern-Simons gauge theory. The main interest in Chern-Simons theory is that it is a topological field theory. In this context, the expectation values of the Wilson loop operators is given in terms of topological invariants associated with the theory of knots, such as their linking numbers and generating functions known as the Jones polynomial. As we will see, the states created by Wilson loops of Chern-Simons theory are anyons, states that obey fractional statistics. Chern-Simons theories are the low energy effective field theories of topological phases such as the topological fluids of the fractional quantum Hall effects.

Chern-Simons gauge theory can be defined on spacetimes of odd dimension D. The simplest case is in $2+1$ dimensions with an abelian gauge group $U(1)$, whose action on a spacetime manifold $\mathcal{M}$ is

$$S = \frac{k}{4\pi}\int^{\mathcal{M}} d^3x\, \epsilon_{\mu\nu\lambda}A^\mu\partial^\nu A^\lambda \equiv \frac{k}{4\pi}\int_{\mathcal{M}} AdA \tag{22.8}$$

where, in the last equality, we introduced the notation of forms.

On a closed manifold $\mathcal{M}$ this action is invariant under both local and large gauge transformations (that wrap around the closed manifold) provided the parameter k, known as the *level*, is an integer. The Chern-Simons action is also odd under time reversal $\mathcal{T}$, $x_0 \to -x_0$, and parity $\mathcal{P}$, defined as $x_1 \to -x_1$ and $x_2 \to x_2$, but it is invariant under $\mathcal{PT}$. The Chern-Simons action is the theory with the smallest number of derivatives that satisfies all these symmetries. For a general non-abelian gauge group G the Chern-Simons action becomes

$$S = \frac{k}{4\pi}\int_{\mathcal{M}} d^3x \; \mathrm{tr}\Big(AdA + \frac{2}{3}A\wedge A\wedge A\Big) \tag{22.9}$$

Here, the cubic term is shorthand for

$$\mathrm{tr}\Big(A\wedge A\wedge A\Big) \equiv \mathrm{tr}\Big(\epsilon_{\mu\nu\lambda}A^\mu A^\nu A^\lambda\Big) \tag{22.10}$$

for a gauge field A^μ that takes values on the algebra of the gauge group G.

A closely related (abelian) gauge theory is the so-called BF theory (Horowitz, 1989) which, in a general spacetime dimension D (even or odd), is a theory of a vector field A^μ (a one-form) and an antisymmetric tensor field B with $D-1$ Lorentz indices (a $D-1$ form), known as a Kalb-Ramond field. Its action is

$$S = \frac{k}{2\pi}\int_{\mathcal{M}} d^Dx \; \epsilon_{\mu\nu\lambda\ldots}B^{\lambda\ldots}\partial^\mu A^\nu \tag{22.11}$$

where, once again, k is an integer. We will see shortly that this theory has the same content as the topological sector of a discrete gauge theory, that we discussed in section 22.2.

Chern-Simons theory was studied by mathematicians in the context of the classification of knots and the representations of the braid group. This theory first entered in physics as the effective action of a theory of N Dirac fields with mass m in 2+1 dimensions, coupled to a background gauge field, and the *parity anomaly* of the fermion determinant.(Deser et al., 1982a; Redlich, 1984a) Up to subtleties related to regularization, the low-energy effective action of the gauge field is found to be (in the abelian case)

$$S_{\text{eff}}[A^\mu] = -i\mathrm{tr}\ln(i\not{D}[A]+m) = \int d^3x \frac{N}{8\pi}\mathrm{sign}(m)AdA + O(1/m) \tag{22.12}$$

and similar expression for the non-abelian case. The fact that, superficially, the level of this effective action is $k = \frac{N}{2}\mathrm{sign}(m)$ requires special care in the regularization of the fermion determinant. We will discuss problem this in the next chapter. This theory has had great success in explaining the phenomenon of statistical transmutation by which bosons turn into fermions by the way of a process of flux attachment. This mechanism is the basis of the modern theory of anyons, particles with fractional statistics (Wilczek, 1982).

22.4 Quantization of Abelian Chern-Simons Gauge Theory

Here we will focus on the simper case of canonical quantization of the abelian Chern-Simons theory in 2+1 dimensions (Dunne et al., 1989; Elitzur et al., 1989; Witten, 1989) Let us consider this theory couple to set of conserved currents j_μ. Its Lagrangian is

$$\mathcal{L} = \frac{k}{4\pi}\epsilon_{\mu\nu\lambda}A^\mu\partial^\nu A^\lambda - j^\mu A_\mu \tag{22.13}$$

which, in Cartesian components it becomes

$$\mathcal{L} = A_0\Big(\frac{k}{4\pi}\epsilon_{ij}\partial_i A_j - j_0\Big) + \frac{k}{4\pi}\epsilon_{ij}A^i\partial^0 A^j + \boldsymbol{j}\cdot\boldsymbol{A} \tag{22.14}$$

As usual, the A_0 component plays the role of a Lagrange multiplier field that enforces a constraint. While in Maxwell's theory the constraint is the Gauss law, $\nabla\cdot\boldsymbol{E} = j_0$, in Chern-Simons theory it implies that

$$\frac{k}{2\pi}B = j_0 \tag{22.15}$$

where $B = \epsilon_{ij}\partial_i A_j$. In other words, this constrain implies a condition that flux and charge must be glued (or attached) to each other. We will see that this constraint is closely related to fractional statistics.

On the other hand, the equation of motion of the gauge field is

$$\frac{k}{2\pi}F^*_\mu = j_\mu \tag{22.16}$$

Therefore, in the absence of external sources, $j_\mu = 0$, the solutions of the equations of motion are the flat connections, $F_{\mu\nu} = 0$, and hence are pure gauge transformations.

At the classical level, the second term of the right hand side of Eq.(22.14) implies that the spatial components of the gauge field, A_i, form canonical pairs. Thus, at the quantum level, as operators they satisfy the equal-time commutation relations

$$\Big[A_1(\boldsymbol{x}), A_2(\boldsymbol{y})\Big] = i\frac{2\pi}{k}\delta(\boldsymbol{x}-\boldsymbol{y}) \tag{22.17}$$

.

Finally, the third term of Eq.(22.14) implies that the Hamiltonian is

$$H = -\int d^2x\, \boldsymbol{j}(\boldsymbol{x})\cdot\boldsymbol{A}(\boldsymbol{x}) \tag{22.18}$$

Hence, in the absence of sources, $\boldsymbol{j} = 0$, the Hamiltonian vanishes, $H = 0$.

The Chern-Simons action is locally gauge-invariant, up to boundary terms.

To see this let us perform a gauge transformation, $A_\mu \to A_\mu + \partial_\mu \Phi$, where $\Phi(x)$ is a smooth, twice differentiable function. Then,

$$S[A^\mu + \partial^\mu \Phi] = \int_{\mathcal{M}} (A^\mu + \partial^\mu \Phi)\epsilon_{\mu\nu\lambda}\partial^\nu (A^\lambda + \partial^\lambda \Phi)$$
$$= \int_{\mathcal{M}} d^3x\ \epsilon_{\mu\nu\lambda} A^\mu \partial^\nu A^\lambda + \int_{\mathcal{M}} d^3x\ \epsilon_{\mu\nu\lambda}\partial^\mu \Phi \partial^\nu A^\lambda \qquad (22.19)$$

Therefore, the change is

$$S[A^\mu + \partial^\mu \Phi] - S[A^\mu] = \int_{\mathcal{M}} d^3x\ \partial^\mu \Phi F^*_\mu$$
$$= \int_{\mathcal{M}} d^3x\ \partial^\mu (\Phi F^*_\mu) - \int_{\mathcal{M}} d^3x\ \Phi \partial^\mu F^*_\mu \qquad (22.20)$$

where $F^*_\mu = \epsilon^{\mu\nu\lambda}\partial_\nu A_\lambda$, is the dual field strength. However, in the absence of magnetic monopoles, this field satisfies the Bianchi identity, $\partial^\mu F^* mu = \partial^\mu(\epsilon_{\mu\nu\lambda}\partial^\nu A^\lambda) = 0$. Therefore, using the Gauss Theorem, we find that the change of the action is a total derivative and integrates to the boundary

$$\delta S = \int_{\mathcal{M}} d^3x\ \partial^\mu (\Phi F^*_\mu) = \int_\Sigma dS_\mu \Phi F^*_\mu \qquad (22.21)$$

where $\Sigma = \partial \mathcal{M}$ is the boundary of $\mathcal{M}$. In particular, if Φ is a non-zero constant function on $\mathcal{M}$, then the change of the action under such a gauge transformation is

$$\delta S = \Phi \times \text{flux}(\Sigma) \qquad (22.22)$$

Hence, the action is not invariant if the manifold has a boundary, and the theory must be supplied with additional degrees of freedom at the boundary.

Indeed, the flat connections, i.e. the solution of the equations of motion, $F_{\mu\nu} = 0$, are pure gauge transformations, $A_\mu = \partial_\mu \varphi$, and have an action that integrates to the boundary. Let the $\mathcal{M} = D \times \mathbb{R}$ where D is a disk in space and $\mathbb{R}$ is time. The boundary manifold is $\Sigma = S^1 \times \mathbb{R}$, where S^1 is a circle. Thus, in this case, the boundary manifold Σ is isotropic to a cylinder. The action of the flat configurations reduces to

$$S = \int_{S^1 \times \mathbb{R}} d^2x \frac{k}{2\pi} \partial_0 \varphi \partial_1 \varphi \qquad (22.23)$$

This implies that the dynamics on the boundary is that of a scalar field on a circle S^1, and obeys periodic boundary conditions.

Although classically the theory does not depend on the metric, it is invariant under arbitrary transformations of the coordinates. However, any gauge fixing condition will automatically break this large symmetry. For instance, we can specify a gauge condition at the boundary in the form of a boundary term of the form $\mathcal{L}_{\text{gauge fixing}} = A_1^2$. In this case, the boundary action of the field φ becomes

$$S[\varphi] = \int_{S^1\times\mathbb{R}} d^2x \, \frac{k}{2\pi} \left[\partial_0\varphi\partial_1\varphi - (\partial_1\varphi^2)\right] \tag{22.24}$$

The solutions of the equations of motion of this compactified scalar field have the form $\varphi(x_1 \mp x_0)$ (where the sign is the sign of k), and are right (left) moving chiral fields depending of the sign of k. This boundary theory is not topological but is is conformally invariant.

A similar result is found in non-abelian Chern-Simons gauge theory. In the case of the $SU(N)_k$ Chern-Simons theory on a manifold $D \times \mathbb{R}$, where D is a disk whose boundary is Γ, and $\mathbb{R}$ is time, the action is

$$S_{\text{CS}}[A] = \int_{D\times\mathbb{R}} d^3x \left[\frac{k}{8\pi}\text{tr}\left(\epsilon_{\mu\nu\lambda}A^\mu\partial^\nu A^\lambda + \frac{2}{3}\epsilon^{\mu\nu\lambda}A_\mu A_\nu A_\lambda\right)\right] \tag{22.25}$$

This theory integrates to the boundary, $\Gamma \times \mathbb{R}$ where it becomes the chiral (right-moving) $SU(N)_k$ Wess-Zumino-Witten model (at level k) at its IR fixed point, $\lambda_c^2 = 4\pi/k$

$$S_{WZW}[g] = \frac{1}{4\lambda_c^2}\int_{\Gamma\times\mathbb{R}} d^2x \; \text{tr}\left(\partial_\mu g\partial^\mu g^{-1}\right) + \frac{k}{12\pi}\int_B \epsilon^{\mu\nu\lambda}\text{tr}\left(g^{-1}\partial_\mu g \; g^{-1}\partial_\nu g \; g^{-1}\partial_\lambda g\right) \tag{22.26}$$

Here, $g \in SU(N)$ parametrizes the flat configurations of the Chern-Simons gauge theory. Therefore, the boundary theory is a non-trivial CFT, the chiral WZW CFT (Witten, 1989).

22.5 Vacuum Degeneracy a Torus

Let us construct the quantum version of this theory on a manifold $\mathcal{M} = T^2 \times \mathbb{R}$, where T^2 is a spatial torus, of linear size L_1 and L_2. Since this manifold does not have boundaries, the flat connections, $\epsilon_{ij}\partial_i A_j = 0$ do not reduce to local gauge transformations of the form $A_i = \partial_i\Phi$. Indeed, the holonomies of the torus T^2, i.e. the Wilson loops on the two non-contractible cycles of the torus Γ_1 and Γ_2 shown in Fig.22.1 are gauge-invariant observables:

$$\int_0^{L_1} dx_1 A_1 \equiv \bar{a}_1, \qquad \int_0^{L_2} dx_1 A_2 \equiv \bar{a}_2 \tag{22.27}$$

where $\bar{a}_1$ and $\bar{a}_2$ are time-dependent. Thus, the flat connections now are

$$A_1 = \partial_1\Phi + \frac{\bar{a}_1}{L_1}, \qquad A_2 = \partial_2\Phi + \frac{\bar{a}_2}{L_2} \tag{22.28}$$

whose action is

$$S = \frac{k}{4\pi} \int dx_0 \epsilon_{ij} \bar{a}_i \partial_0 \bar{a}_j \tag{22.29}$$

Therefore, the global degrees of freedom $\bar{a}_1$ and $\bar{a}_2$ at the qunatum level become operators that satisfy the commutation relations

$$[\bar{a}_1, \bar{a}_2] = i\frac{2\pi}{k} \tag{22.30}$$

We find that the flat connections are described by the quantum mechanics of $\bar{a}_1$ and $\bar{a}_2$. A representation of this algebra is

$$\bar{a}_2 \equiv -i\frac{2\pi}{k}\frac{\partial}{\partial \bar{a}_1} \tag{22.31}$$

Furthermore, the Wilson loops on the two cycles become

$$W[\Gamma_1] = \exp\left(i \int_0^{L_1} A_1\right) \equiv e^{i\bar{a}_1}, \quad W[\Gamma_2] = \exp\left(i \int_0^{L_2} A_2\right) \equiv e^{i\bar{a}_2} \tag{22.32}$$

and satisfy the algebra

$$W[\Gamma_1]W[\Gamma_2] = \exp(-i2\pi/k)W[\Gamma_2]W[\Gamma_1] \tag{22.33}$$

Under large gauge transformations

$$\bar{a}_1 \to \bar{a}_1 + 2\pi, \qquad \bar{a}_2 \to \bar{a}_2 + 2\pi \tag{22.34}$$

Therefore, invariance under large gauge transformations on the torus implies that $\bar{a}_1$ and $\bar{a}_2$ define a two-torus target space.

Let us define the unitary operators

$$U_1 = \exp(ik\bar{a}_2), \quad U_2 = \exp(-ik\bar{a}_1) \tag{22.35}$$

which satisfy the algebra

$$U_1 U_2 = \exp(i2\pi k) U_2 U_1 \tag{22.36}$$

The unitary transformations U_1 and U_2 act as shift operators on $\bar{a}_1$ and $\bar{a}_2$ by 2π, and hence generate the large gauge transformations. Moreover, the unitary operators U_1 and U_2 leave the Wilson loop operators on non-contractible cycles invariant,

$$U_1^{-1} W[\Gamma_1] U_1 = W[\Gamma_1], \quad U_2^{-1} W[\Gamma_2] U_2 = W[\Gamma_2] \tag{22.37}$$

Let $|0\rangle$ be the eigenstate of $W[\Gamma_1]$ with eigenvalue 1, i.e. $W[\Gamma_1]|0\rangle = |0\rangle$. The state $W[\Gamma_2]|0\rangle$ is also an eigenstate of $W[\Gamma_1]$ with eigenvalue $\exp(-i2\pi/k)$, since

$$W[\Gamma_1]W[\Gamma_2]|0\rangle = e^{i2\pi/k} W[\Gamma_2]W[\Gamma_1]|0\rangle = e^{-i2\pi/k} W[\gamma_2]|0\rangle \tag{22.38}$$

More generally, since

$$W[\Gamma_1]W^p[\Gamma_2]|0\rangle = e^{-i2\pi p/k}W^p[\Gamma_2]|0\rangle \tag{22.39}$$

we find that, provided $k \in \mathbb{Z}$, there are k linearly independent vacuum states $|p\rangle = W^p[\Gamma_2]|0\rangle$, for the $U(1)$ Chern-Simons gauge theory at level k. It is denoted as the $U(1)_k$ Chern-Simons theory. Therefore the finite-dimensional topological space on a two-torus is k-dimensional. It is trivial to show that, on a surface of genus g, the degeneracy is k^g.

We see that in the abelian $U(1)_k$ Chern-Simons theory the Wilson loops must carry k possible values of the unit charge. This property generalizes to the non-abelian theories, where it technically more subtle. We will only state some important results. For example, if the gauge group is $SU(2)$ we expect that the Wilson loops will carry the representation labels of the group $SU(2)$, i.e. they will be labelled by (j, m), where $j = 0, \frac{1}{2}, 1, \ldots$ and the $2j+1$ values of m satisfy $|m| \leq j$. However, it turns out $SU(2)_k$ Chern-Simons theory has fewer states, and that the values of j are restricted to the range $j = 0, \frac{1}{2}, \ldots, \frac{k}{2}$.

22.6 Fractional Statistics

Another aspect of the topological nature of Chern-Simons theory is the behavior of expectation values of products of Wilson loop operators. Let us compute the expectation value of a product of two Wilson loop operators on two positively oriented closed contours γ_1 and γ_2. We will do this computation in the abelian Chern-Simons theory $U(1)_k$ in 2+1-dimensional Euclidean space. Note that the Euclidean Chern-Simons action is pure imaginary since the action is first-order in derivatives. The expectation value to be computed is

$$W[\gamma_1 \cup \gamma_2] = \left\langle \exp\left(i \oint_{\gamma_1 \cup \gamma_2} dx_\mu A_\mu\right)\right\rangle_{\text{CS}} \tag{22.40}$$

We will see that the result changes depending on whether the loops γ_1 and γ_2 are linked or unlinked, as in the cases shown in Fig.22.3a and b.

This calculation is simpler than the one we did for Maxwell's theory in Section 9.7. As in Maxwell's case, the expectation value of a Wilson loop on a contour (or union of contours, as in the present case) γ can be written as

$$\left\langle \exp\left(i \oint_\gamma dx_\mu A_\mu\right)\right\rangle_{\text{CS}} = \left\langle \exp\left(i \int d^3x J_\mu A_\mu\right)\right\rangle_{\text{CS}} \tag{22.41}$$

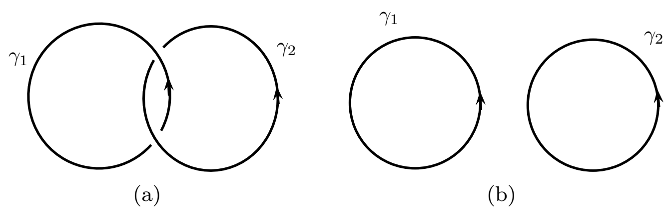

Figure 22.3 a) Two linked Wilson loops forming a knot; b) two unliked Wilson loops.

where the current J_μ is

$$J_\mu(x) = \delta(x_\mu - z_\mu(t))\frac{dz_\mu}{dt} \tag{22.42}$$

Here $z_\mu(t)$ is a parametrization of the contour γ. Therefore, the expectation value of the Wilson loop is (Witten, 1989)

$$\begin{aligned}\left\langle \exp\left(i\oint_\gamma dx_\mu A_\mu\right)\right\rangle_{\text{CS}} &\equiv \exp(iI[\gamma]_{\text{CS}})\\ &= \exp\left(-\frac{i}{2}\int d^3x \int d^3y\ J_\mu(x)G_{\mu\nu}(x-y)J_\nu(y)\right)\end{aligned} \tag{22.43}$$

where $G_{\mu\nu}(x-y) = \langle A_\mu(x)A_\nu(y)\rangle_{\text{CS}}$ is the propagator of the Chern-Simons gauge field. Since the loops are closed, the current J_μ is conserved, i.e. $\partial_\mu J_\mu = 0$, and the effective action $I[\gamma]_{\text{CS}}$ of the loop γ is gauge-invariant.

The Euclidean propagator of Chern-Simons gauge theory (in the Feynman gauge) is

$$G_{\mu\nu}(x-y) = \frac{2\pi}{k}G_0(x-y)\epsilon_{\mu\nu\lambda}\partial_\lambda\delta(x-y) \tag{22.44}$$

where $G_0(x-y)$ is the propagator of the massless Euclidean scalar field, which satisfies

$$-\partial^2 G_0(x-y) = \delta^3(x-y) \tag{22.45}$$

As usual, we can write

$$G_0(x-y) = \langle x|\frac{1}{-\partial^2}|y\rangle \tag{22.46}$$

Using these results, we find the following expression for the effective action

$$
\begin{aligned}
I[\gamma]_{\rm CS} =&\frac{\pi}{k}\int d^3x \int d^3y\; J_\mu(x)J_\nu(y)G_0(x-y)\epsilon_{\mu\nu\lambda}\partial_\lambda\delta(x-y)\\
=&\frac{\pi}{k}\oint_\gamma dx_\mu \oint_\gamma dy_\nu \epsilon_{\mu\nu\lambda}\partial_\lambda G_0(x-y)
\end{aligned}
\tag{22.47}
$$

Again, since the current J_μ is conserved, it can be written as the curl of a vector field, B_μ, as

$$J_\mu = \epsilon_{\mu\nu\lambda}\partial_\nu B_\lambda \tag{22.48}$$

In the Lorentz gauge, $\partial_\mu B_\mu = 0$, we can write

$$B_\mu = \epsilon_{\mu\nu\lambda}\partial_\nu \phi_\lambda \tag{22.49}$$

Hence,

$$J_\mu = -\partial^2 \phi_\mu \tag{22.50}$$

where

$$\phi_\mu(x) = \int d^3y\; G_0(x-y)J_\mu(y) \tag{22.51}$$

Upon substituting this result into the expression for B_μ, we find

$$B_\mu = \int d^3y \epsilon_{\mu\nu\lambda}\partial_\nu G_0(x-y)J_\lambda(y) = \oint_\gamma \epsilon_{\mu\nu\lambda}\partial_\nu G_0(x-y)dy_\lambda \tag{22.52}$$

Therefore, the effective action $I[\gamma]_{\rm CS}$ becomes

$$I[\gamma]_{\rm CS} = \frac{\pi}{k}\oint_\gamma dx_\mu \oint_\gamma dy_\nu \epsilon_{\mu\nu\lambda}\partial_\lambda G_0(x-y) = \frac{\pi}{k}\oint_\gamma dx_\mu B_\mu(x) \tag{22.53}$$

Let Σ be an oriented open surface of the Euclidean three-dimensional space whose boundary is the oriented loop (or union of loops) γ, i.e. $\partial\Sigma = \gamma$. Then, using Stokes Theorem we write in the last line of Eq.(22.53) as

$$I[\gamma]_{\rm CS} = \frac{\pi}{k}\int_\Sigma dS_\mu \epsilon_{\mu\nu\lambda}\partial_\nu B_\lambda = \frac{\pi}{k}\int_\Sigma dS_\mu J_\mu \tag{22.54}$$

The integral in the last line of this equation cis the flux of the current J_μ through the surface Σ. Therefore, this integral counts the number of times n_γ the Wilson loop on γ pierces the surface Σ (whose boundary is γ), and therefore it is an integer, $n_\gamma \in \mathbb{Z}$. We will call this integer the *linking number* (or Gauss invariant) of the configuration of loops. In other words, the expectation value of the Wilson loop operator is

$$W[\gamma]_{\rm CS} = \exp\left(i\frac{\pi}{k}n_\gamma\right) \tag{22.55}$$

The linking number is a *topological invariant* since, being an integer, its value cannot be changed by smooth deformations of the loops, provided they are not allowed to cross.

We will now show that this property of Wilson loops in Chern-Simons gauge theory leads to the concept of fractional statistics. Let us consider a scalar matter field that is massive and charged under the Chern-Simons gauge field. The excitations of this matter field are particles that couple minimally to the gauge field. Here we will be interested in the case in which these particles are very heavy. In that limit, we can focus on states that have a few of this particles which will be in their non-relativistic regime.

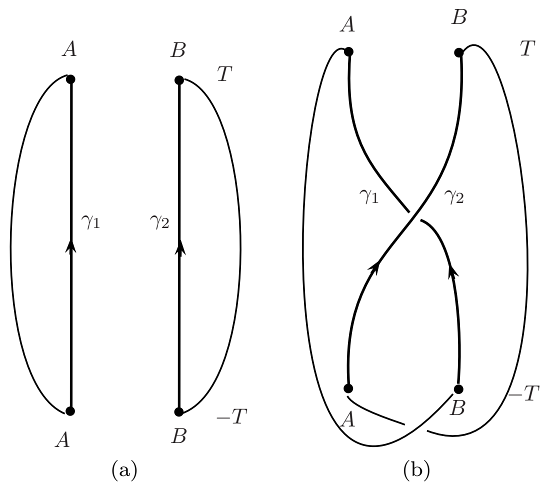

Figure 22.4 a) A topological link of two Wilson loops forming a knot; b) two unliked Wilson loops.

Consider, for example, a state with two particles which in the remote past, at time $t = -T \to -\infty$, are located at two points A and B. This initial state will evolve to a final state at time $t = T \to \infty$, in which the particles either go back to their initial locations (the direct process), or to another one in which they exchange places, $A \leftrightarrow B$. At intermediate times, the particles follow smooth worldlines. These two processes, direct and exchange, are shown in Figs.22.4 a and b. There we see that the direct process is equivalent to a history with two unlinked loops (the worldlines of the particles), whereas in the exchange process the two loops form a link. It follows from the preceding discussion that the two amplitudes differ by the result of the computation of the Wilson loop expectation value for the loops γ_1 and γ_2. Let us call the first amplitude W_{direct} and the second W_{exchange}. The result is

$$W_{\text{exchange}} = W_{\text{direct}} \exp(\pm i\pi/k) \tag{22.56}$$

where the sign depends on how the two worldlines wind around each other.

An equivalent interpretation of this result is that if $\Psi[A, B]$ is the wave function with the two particles at locations A and B, the wavefunction where their locations are exchanged is

$$\Psi[B, A] = \exp(\pm i\pi/k)\Psi[A, B] \tag{22.57}$$

Clearly, for $k = 1$ the wave function is antisymmetric and the particles are fermions, while for $k \to \infty$ they are bosons. At other values of k the particles obey *fractional statistics* and are called *anyons*. The phase factor $\phi = \pm\pi/k$ is called the statistical phase.

Notice that, while for fermions and bosons the statistical phase $\varphi = 0, \pi$ is uniquely defined (mod 2π), for other values of k the statistical angle is specified up to a sign that specifies how the worldlines wind around each other. Indeed, mathematically the exchange process shown in Fig.22.4b is known as a *braid*. Processes in which the worldlines wind clock and counterclockwise are braids that are inverse of each other. Braids can also be stack on top of each other yielding multiples of the phase φ. In addition to stacking braids, Wilson loops can be fused: seen from some distance, a pair of particles will behave as a new particle with a well defined behavior under braiding. This process of fusion is closely related to the concept of fusion of primary fields in Conformal Field Theory, discussed in Chapter 21.

What we have just described is a mathematical structure called the *Braid* group. The example that we worked out using abelian Chern-Simons theory yields one-dimensional representations of the Braid group with the phase φ being the label of the representations. For $U(1)_k$ there are k types of particles (anyons). That these representations are abelian means that, in the general case of $U(1)_k$, acting on a one-dimensional representation p (defined mod k) with a one-dimensional representation q (also defined mod k) yields the representation one-dimensional $p + q$ (mod k). We will denote the operation of *fusing* these representations (particles!) as $[q]_{\text{mod } k} \times [p]_{\text{mod } k} = [q + p]_{\text{mod } k}$. These representations are in one-to-one correspondence with the inequivalent charges of the Wilson loops, and with the vacuum degeneracy of the $U(1)_k$ Chern-Simons theory on a torus.

A richer structure arises in the case of the non-abelian Chern-Simons theory at level k (Witten, 1989), such as $SU(2)_k$. For example, for $SU(2)_1$ the theory has only two representations, both are one-dimensional, and have statistical angles $\varphi = 0, \pi/2$.

However, for $SU(2)_k$, the content is more complex. In the case of $SU(2)_2$ the theory has a) a trivial representation $[0]$ (the identity, $(j,m) = (0,0)$), b) a (spinor) representation $[1/2]$ ($(j,m) = (1/2, \pm 1/2)$), and c) a the representation $[1]$ ($(j,m) = (1,m)$, with $m = 0, \pm 1$). These states will fuse obeying the following rules: $[0] \times [0] = [0]$, $[0] \times [1/2] = [1/2]$, $[0] \times [1] = [1]$, $[1/2] \times [1/2] = [0] + [1]$, $[1/2] \times [1] = [1/2]$, and $[1] \times [1] = [0]$ (note the truncation of the fusion process!).

Of particular interest is the case $[1/2] \times [1/2] = [0] + [1]$. In this case we have two fusion channels, labeled by $[0]$ and $[1]$. The braiding operations now will act on a two-dimensional Hilbert space and are represented by 2×2 matrices. This is an example of a non-abelian representation of the braid group. These rather abstract concepts have found a physical manifestation in the physics of the fractional quantum Hall fluids, whose excitations are vortices that carry fractional charge and anyon (braid) fractional statistics.

Why this is interesting can be seen by considering a Chern-Simons gauge theory with four quasi-static Wilson loops. For instance in the case of the $SU(2)_2$ Chern-Simons theory the Wilson loops (heavy particles!) can be taken to carry the spinor representation, $[1/2]$. If we call the four particles A, B, C and D, we would expect that their quantum state would be completely determined by the coordinates of the particles. This, however, is not the case since, if we fuse A with B, the result is either a state $[0]$ or a state $[1]$. Thus, if the particles were prepared originally in some state, braiding (and fusion) will lead to a linear superposition of the two states. This braiding process defines a unitary matrix, a representation of the Braid Group. The same is true with the other particles. However, it turns out that for four particles there only two linearly independent states. This two-fold degenerate Hilbert space of topological origin is called a topological qubit. Moreover, if we consider a system with N (even) number of such particles, the dimension of the topologically protected Hilbert space is $2^{\frac{N}{2}-1}$. Hence, for large N, the entropy per particle grows as $\frac{1}{2} \ln 2 = \ln \sqrt{2}$. Therefore the qubit is not an "internal" degree of freedom of the particles but a collective state of topological origin. Interestingly, there are physical systems, known as non-abelian fractional quantum Hall fluids that embody this physics and are accessible to experiments! For these reasons, the non-abelian case has been proposed as a realization of a topological qubit (Kitaev, 2003; Das Sarma et al., 2008).

We close with a few comments on BF theories in 2+1 dimensions. With minor changes, the statements below generalize to higher dimensions. These theories are abelian and in 2+1 dimensions the to fields are one-forms, see

Eq.(22.11). The BF lagrangian implies that the spatial components two fields, that we denoted by A_μ and B_μ, form canonical pairs, and satisfy the equal-time commutations relations

$$[A_i(\boldsymbol{x}), B_j(\boldsymbol{y})] = i\frac{\pi}{k}\epsilon_{ij}\delta(\boldsymbol{x}-\boldsymbol{y}) \tag{22.58}$$

As a result, on a two-torus T^2, the Wilson loop of the A_μ field on the non-contractible loop Γ_i, $W^A[\Gamma_i]$, and the Wilson loop of the B_μ field on the non-contractible loop Γ_j, $W^B[\Gamma_j]$, do not commute, and obey the algebra

$$W^A[\Gamma_i]W^B[\Gamma_j] = \epsilon_{ij}\exp(\pm i\pi/k)W^B[\Gamma_j]W^A[\Gamma_i] \tag{22.59}$$

We recognize that, for the case $k = 1$, this is the same algebra obeyed by the Wilson and 't Hooft loops on two non-contractible cycles of T^2 of $\mathbb{Z}_2$ gauge theory in its deconfined phase. It is easy to show that, in fact, the content of the BF at level k is the same as that of the $\mathbb{Z}_k$ gauge theory in its deconfined phase, and that the BF theory at level k is an equivalent (continuum) representation.