

Fractionalization, Topological Order, and Quasiparticle Statistics

Masaki Oshikawa¹ and T. Senthil^{2,3}

¹*Department of Physics, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo 152-8551, Japan*

²*Center for Condensed Matter Theory, Department of Physics, Indian Institute of Science, Bangalore 560 012, India*

³*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 13 May 2005; published 13 February 2006)

We argue, based on general principles, that topological order is essential to realize fractionalization in gapped insulating phases in dimensions $d \geq 2$. In $d = 2$ with genus g , we derive the existence of the minimum topological degeneracy q^g if the charge is fractionalized in units of $1/q$, irrespective of microscopic model or effective theory. Furthermore, if the quasiparticle is either boson or fermion, it must be at least q^{2g} .

DOI: 10.1103/PhysRevLett.96.060601

PACS numbers: 05.30.Pr, 71.10.Hf, 71.10.Pm, 73.43.Cd

Fractionalization of quantum numbers has been a focus of condensed matter physics in recent years. It refers to the emergence of a collective excitation having fractional quantum numbers with respect to the elementary particles (such as electrons), in a strongly correlated system. The notion of fractionalization is not only fascinating in itself, but also has been related to other intriguing concepts in theoretical physics as discussed in the following.

At present, several different systems exhibit the fractionalization [1–10], at least theoretically. While the details naturally depend on each model under consideration, the structure of the excitation spectrum is efficiently described in terms of a gauge theory. More precisely the excitations consist of objects that have long ranged nonlocal “statistical” interactions with each other, which may be encoded as an Aharonov-Bohm gauge interaction. This is well known in the fractional quantum hall effect where the fractionalized quasiparticles also have fractional statistics. Similarly in the fractionalized liquids described in Refs. [3–8], there are vison excitations that have long range statistical interactions with the fractionalized particles (such as the spinons in a spin liquid).

This emergent gauge structure generally implies the existence of a certain kind of order—dubbed topological order—associated with the global properties of the ground-state wave function [11], which is also commonly found in the above examples. A characteristic signature of the topological order is the ground-state degeneracy depending on the topology of the system. This cannot be understood as a consequence of a conventional spontaneous symmetry breaking, which is the standard mechanism behind the ground-state degeneracy. The intriguing nature and consequences of the topological order is best understood in the fractional quantum Hall liquid (FQHL) [11], although some of the concepts are applicable to other systems. In the gauge theory picture, the topological degeneracy could be understood with different “vacua” corresponding to a different number of vortices trapped in each “hole” of the space (such as the torus).

However, these developments apparently leave open the question on whether there is a different way to realize

fractionalization without the emergent gauge structures. As such, at this point it is also unclear whether the topological order and associated ground-state degeneracy are necessary to have fractionalization. As we are still far from the complete classification of the fractionalized phases, and many novel examples of fractionalization will likely be found in the future, these questions are of significant importance.

In this Letter, we demonstrate that there is indeed a general and direct connection between the fractionalization and the topological order, in the specific context of systems with a fully gapped spectrum. Generalizing the gauge invariance argument presented in Ref. [12], the existence of the topological order is shown to follow just from the fractionalization, irrespective of microscopic details.

As discussed before, the known examples of the fractionalization are rather suggestive of such a universal relation. In the several examples of fractionalization (at zero magnetic field) discussed recently [3–10], the degeneracy is at least q^{2g} -fold in a system on a $d = 2$ dimensional surface with genus g , if the fractionalization occurs in the unit of $1/q$. However, we must recall that the Laughlin FQHL does exhibit a ground-state degeneracy, but only q^g -fold [13,14]. (See also Ref. [10].) We will also clarify the difference between the cases with q^{2g} -fold and q^g -fold degeneracy, which turns out to be related to the statistics of the quasiparticles. Not surprisingly, our argument is closely related to the earlier studies on the topological order in FQHL especially in Refs. [13,14], and that in systems of anyons [15].

Now let us define the problem in a general setting. We consider a system defined by a certain microscopic Hamiltonian of interacting particles, with an exact $U(1)$ global symmetry. With the global $U(1)$ symmetry, we may assign a (fictitious) charge to each particle, with the total charge being a conserved quantity. We take the unit in which the elementary charge is unity, so that the charge of all the particles appearing in the microscopic model is an integer. We can now also introduce a (fictitious) external $U(1)$ gauge field (“electromagnetic field”) coupled to the

charge. We set $\hbar = c = 1$ so that the unit flux quantum is given by 2π .

The ground state is generally a complicated state in terms of the original particles. Here we assume for simplicity that there is a finite gap above the (possibly degenerate) ground state(s). We further assume that the elementary excitations of the system are well-defined quasiparticles and quasiholes. The quasiparticle may carry a charge that is a fraction of the original unit charge, *thereby we define the fractionalization*. This definition is very natural and is independent of the concrete model or mechanism of the fractionalization, while it naturally applies to all the known cases. Let us assume that the fractional charge of the quasiparticle is p/q , where p and q are mutually prime integers.

For simplicity, let us, for a moment, consider a system on a $d = 2$ torus of sufficiently large size $L_x \times L_y$. We will comment on other cases later. We define the following process as introduced in Ref. [14]. First, we create a quasiparticle and its antiparticle (quasihole) out of the vacuum (ground state) at some location, and then move the quasiparticle to the $+x$ direction, so that it encircles the torus to come back to the original location and to meet the quasihole. Finally, we pair annihilate the quasiparticle and quasihole. Here we assume that this process can be realized by a unitary time-evolution operator \mathcal{T}_x with respect to a properly chosen time-dependent Hamiltonian, e.g., with a time-dependent local potential to create and drag the quasiparticle. Thus we exclude quantum “glassy” systems as proposed in Ref. [16].

Similarly, we introduce another unitary operator \mathcal{T}_y , corresponding to creation of a quasihole-quasiparticle pair and annihilation after winding in the y direction. It is expected that $\mathcal{T}_{x,y}$ bring any state in the ground-state manifold state back, at least approximately, to a (possibly different) ground state.

Next we consider an adiabatic insertion of a unit flux quantum $\Phi_0 = 2\pi$ through the hole of the torus, inducing an (fictitious) electric field in the x direction. This may again be realized by a time evolution in which the x component of the vector potential is gradually increased from $A_x = 0$ to $A_x = 2\pi/L_x$ in the Hamiltonian. Thus it is represented by a unitary time-evolution operator \mathcal{F}_x . We also define a similar operator \mathcal{F}_y that corresponds to an adiabatic insertion of a unit flux quantum through the other hole, inducing the y component of the vector potential. We assume that these operations do not close the gap to excitations above the ground-state manifold and thus bring any ground state to a ground state. In $d = 2$, it amounts to assuming the system to be an insulator, while it is a stronger assumption for $d \geq 3$ [17]. (See also Ref. [18].)

Now let us consider two operations \mathcal{T}_x and \mathcal{F}_x in sequence. The flux insertion \mathcal{F}_x introduces the vector potential $A_x = 2\pi/L_x$, corresponding to the unit flux quantum $\Phi_0 = 2\pi$ contained in the system. As we consider the

process \mathcal{T}_x in different backgrounds, let us distinguish them by denoting $\mathcal{T}_x(\Phi)$ as the “encircling” process defined above in the presence of the vector potential $A_x = \Phi/L_x$. The contained unit flux quantum does not induce the Aharonov-Bohm effect on the original particles of integral charge. However, for the quasiparticle with the fractional charge p/q , the same vector potential $A_x = 2\pi/L_x$ still gives a nontrivial Aharonov-Bohm phase $e^{2\pi ip/q}$ when the quasiparticle completes the encircling process. Thus we obtain a relation

$$\mathcal{T}_x(\Phi_0)\mathcal{F}_x \sim e^{2\pi ip/q} \mathcal{F}_x \mathcal{T}_x(0). \quad (1)$$

On the other hand, because the microscopic model is given in terms of the original particles of integral charge, any Hamiltonian with an extra unit flux quantum in the hole of the torus is unitary equivalent to the Hamiltonian with zero flux. Namely, the Hamiltonian $\mathcal{H}(\Phi_0)$ with the unit flux quantum is related to the Hamiltonian $\mathcal{H}(0)$ with zero flux as $\mathcal{H}(\Phi_0) = U^{-1} \mathcal{H}(0)U$, by a unitary operator U which is called as the large gauge transformation. As we have argued previously, the encircling process \mathcal{T}_x should be realized by a time evolution with respect to an appropriately chosen time-dependent Hamiltonian, again *written in terms of the original particles*. Therefore, the operator \mathcal{T} also must obey the relation

$$\mathcal{T}_x(\Phi_0) = U^{-1} \mathcal{T}_x(0)U. \quad (2)$$

Combining Eqs. (1) and (2), we obtain

$$\mathcal{T}_x(0)\tilde{\mathcal{F}}_x \sim e^{2\pi ip/q} \tilde{\mathcal{F}}_x \mathcal{T}_x(0), \quad (3)$$

where $\tilde{\mathcal{F}}_x \equiv U\mathcal{F}_x$. In the following, for brevity \mathcal{T}_x without the argument denotes $\mathcal{T}_x(0)$, and likewise for \mathcal{T}_y .

This algebra between \mathcal{T}_x and $\tilde{\mathcal{F}}_x$ is identical to that of the magnetic translation group, which we call as the magnetic algebra. By our assumptions, both of these operators map a ground state to a ground state. We thus immediately see that the ground states must be q -fold degenerate, with the same reasoning as was used in Refs. [14,15]. Our argument so far is essentially contained in Ref. [12], where the q -fold ground-state degeneracy of a FQHL is derived based on the gauge invariance. In this Letter, we shall present a more systematic discussion to demonstrate that the degeneracy is topology dependent, and that the degeneracy is also affected by quasiparticle statistics.

For the other direction y , we obtain a corresponding relation

$$\mathcal{T}_y \tilde{\mathcal{F}}_y \sim e^{2\pi ip/q} \tilde{\mathcal{F}}_y \mathcal{T}_y. \quad (4)$$

Apparently, now we obtain two sets of the magnetic algebra, which would imply a q^2 -fold degeneracy on the torus. However, as it should not apply to the Laughlin state where the degeneracy is known to be only q -fold, we have to examine more carefully the interplay between Eqs. (3) and (4).

\mathcal{F}_x introduces the vector potential only in the x direction, to which \mathcal{T}_y is insensitive. Thus, with the large gauge transformation combined, we have $\tilde{\mathcal{F}}_x \mathcal{T}_y = \mathcal{T}_y \tilde{\mathcal{F}}_x$ and likewise for $\tilde{\mathcal{F}}_y$ and \mathcal{T}_x . Therefore, we can take the basis in the ground-state subspace so that $\tilde{\mathcal{F}}_x$ and \mathcal{T}_y are both diagonalized. Let the simultaneous eigenstate (among the ground states) of them be $|f_x, t_y\rangle$, with f_x and t_y denoting the eigenvalues of $\tilde{\mathcal{F}}_x$ and \mathcal{T}_y , respectively. By applying \mathcal{T}_x to this state, one obtains a new ground state belonging to a different eigenvalue $f_x e^{-2\pi i p/q}$ of $\tilde{\mathcal{F}}_x$ because

$$\tilde{\mathcal{F}}_x(\mathcal{T}_x|f_x, t_y\rangle) = f_x e^{-2\pi i p/q}(\mathcal{T}_x|f_x, t_y\rangle) \quad (5)$$

follows from Eq. (3). By repeated applications of \mathcal{T}_x , one can obtain at least q different ground states as announced.

Similarly, we can apply $\tilde{\mathcal{F}}_y$ to $|f_x, t_y\rangle$ to obtain q -fold degenerate ground states belonging to different eigenvalues of \mathcal{T}_y . The question now is whether these two procedures give different set of ground states. It depends on whether (or how) the application of \mathcal{T}_x changes the eigenvalue of \mathcal{T}_y . This boils down to the commutation relation between \mathcal{T}_x and \mathcal{T}_y , which actually reflects the statistics of the quasiparticle. The (Abelian) anyonic statistic is characterized by a statistical angle θ , so that an exchange of two identical particles gives rise to the phase factor $e^{-i\theta}$. In Refs. [14,15] it was pointed out that

$$\mathcal{T}_x^{-1} \mathcal{T}_y^{-1} \mathcal{T}_x \mathcal{T}_y = e^{-i2\theta}. \quad (6)$$

This is because the left-hand side corresponds to world lines of the two quasiparticles forming two linked loops in the spacetime, as illustrated in Figs. 4 and 5 in Ref. [14].

Let us first consider the simple case of either bosonic ($\theta = 0$) or fermionic ($\theta = \pi$) statistics, for which \mathcal{T}_x and \mathcal{T}_y commute from Eq. (6). Thus, applying \mathcal{T}_x does not change the eigenvalue t_y of \mathcal{T}_y while it changes the eigenvalue of $\tilde{\mathcal{F}}_x$. Therefore, in this case, one can obtain q different eigenvalues of $\tilde{\mathcal{F}}_x$ by successively applying \mathcal{T}_x , for each of q different eigenvalues of \mathcal{T}_y that are obtained by application of $\tilde{\mathcal{F}}_y$. Thus, there are at least q^2 -fold degenerate ground states corresponding to the different set of eigenvalues. In particular, when $\tilde{\mathcal{F}}_x$ and $\tilde{\mathcal{F}}_y$ commute, the degeneracy deduced from the above set of algebra is q^2 .

On the other hand, in the Laughlin state at filling fraction $1/q$ where q is odd, the quasiparticles are known to carry the fractional charge $1/q$ ($p = 1$ in the previous notation) [1], and to exhibit anyonic fractional statistics with the statistical angle $\theta = \pi/q$ [19]. In this case, because of Eq. (6), we obtain

$$\mathcal{T}_y(\mathcal{T}_x|f_x, t_y\rangle) = t_y e^{2\pi i/q}(\mathcal{T}_x|f_x, t_y\rangle). \quad (7)$$

Thus, combined with Eq. (5), an application of \mathcal{T}_x induces the change in both the eigenvalues of $\tilde{\mathcal{F}}_x$ and \mathcal{T}_y as

$$(f_x, t_y) \rightarrow (f_x e^{-2\pi i/q}, t_y e^{2\pi i/q}). \quad (8)$$

This allows the possibility that the ground-state degeneracy on the torus to be smaller than q^2 . This could happen if

$$\tilde{\mathcal{F}}_x \tilde{\mathcal{F}}_y \sim e^{-2\pi i/q} \tilde{\mathcal{F}}_y \tilde{\mathcal{F}}_x, \quad (9)$$

when acting on the ground-state subspace. In this case, because the application of $\tilde{\mathcal{F}}_y$ induces exactly the same change of the eigenvalues in Eq. (8), we can generate only q different set of eigenvalues.

In fact, Eq. (9) is exactly what holds in the Laughlin state. As pointed out in Ref. [14], because the quasiparticles and holes in the Laughlin state can be identified with a “vortex” with unit flux quantum, the encircling process \mathcal{T}_x actually introduces a unit flux quantum threading the hole of the torus, as \mathcal{F}_y does. Thus Eq. (9) follows. Actually, it means that $\mathcal{T}_{x,y}$ can be identified with $\mathcal{F}_{y,x}$ as far as their action in the ground-state subspace is concerned. Thus the two algebra Eqs. (3) and (4) are indeed reduced to a single magnetic algebra, leaving only the q -fold degeneracy. On the other hand, if the statistical angle θ does not match the fractional charge of the quasiparticles, we should have a larger degeneracy. When the quasiparticle statistics is non-Abelian, the exact counting is more complicated. Nevertheless, the minimum q -fold degeneracy still holds because Eq. (3) is based on the fractionalized charge and should not depend on the statistics. The detailed discussion of the non-Abelian case is deferred to a separate publication.

The above discussion can be generalized to a two-dimensional system on the surface with genus g , for which there are g pairs of intersecting elementary nontrivial cycles. We can define the flux insertion (plus the appropriate large gauge transformation) operator $\tilde{\mathcal{F}}_c$ and the quasiparticle winding operator \mathcal{T}_c for each cycle c . Picking one cycle from each pair, we have a set of g nonintersecting cycles so that the operators for the different cycles commute. Thus, for any (Abelian or non-Abelian) statistics of the quasiparticles, we have g independent magnetic algebras acting on the ground-state subspace and thus the ground-state degeneracy must be at least q^g . If the quasiparticle is either a boson or fermion, we can utilize a set of $2g$ magnetic algebras and the degeneracy must be at least q^{2g} .

The close relation between the insertion of the unit flux quantum and trapped vortices was emphasized previously in the \mathbf{Z}_2 gauge theory description of a fractionalized phase [4]. The adiabatic flux insertion $\tilde{\mathcal{F}}_c$ was also used to relate topologically degenerate ground states in the FQHL [11,14]. The present argument suggests that these structures are rather universal in fractionalized systems.

Our argument could also be generalized to dimensions $d \neq 2$. Although our understanding of the topological order is still incomplete for $d \geq 3$, our argument implies a ground-state degeneracy in a gapped fractionalized sys-

tem defined on a geometry with a nontrivial fundamental group. This suggests that the topological order is essential for the fractionalization, also in $d \geq 3$.

On the other hand, the situation is quite different in $d = 1$, where the “polyacetylene”-type fractionalization [20] is known to occur in a conventional ordered phase with a spontaneous breaking of the translation symmetry. Our argument applied to $d = 1$ just requires the ground state on a ring to be degenerate, as there is no higher topology. The degeneracy can be understood as a consequence of the spontaneous symmetry breaking of the conventional type, rather than due to any “topological” order. For $d \geq 2$, our argument reveals that the ground-state degeneracy indeed depends on the topology, implying the topological order. The present observation could help clarifying the profound difference in the fractionalization between $d = 1$ and $d \geq 2$.

Throughout this Letter we have assumed the system to have a finite gap. However, the topological order can exist also in gapless systems [4], which we have not yet analyzed. It might be interesting to extend our argument to gapless cases. Although the concept of the ground-state degeneracy itself becomes subtle, the (quasi-)degenerate ground states may be identified separately from the gapless excitations, for example, by examining the finite-size scaling carefully.

To summarize, we have derived a topological degeneracy, which indicates the presence of a topological order, in a general (gapful) fractionalized system in $d = 2$. The magnitude of the degeneracy is also related to the statistics of the fractionalized quasiparticles. It is also notable that the simple trick of the flux insertion together with the gauge invariance leads to the rather strong statement, to be added to existing applications [13,14,17,18,21–24]. Comparing with the “momentum counting”-type applications [13,17,18,22–24] of the flux insertion, the present argument is more powerful in the sense that it can be applied to various topologies, to show the degeneracy is indeed topological. The momentum counting arguments can be applied only to a cylinder or a torus, and thus by itself does not indicate whether the derived degeneracy is a topological one or due to a conventional order. On the other hand, the operator \mathcal{T}_c needed in the present argument is introduced in a hand-waving way, and thus makes the argument considerably less rigorous than the momentum counting ones.

This work was initiated during a Summer Workshop at Aspen Center for Physics in which both of the authors participated, and extended while M. O. attended a program at Kavli Institute for Theoretical Physics, UC Santa Barbara (supported in part by NSF Grant No. PHY99-

07949). M. O. thanks the participants of the program at KITP, and Claudio Chamon for very useful comments. In particular, we are grateful to Gregoire Misguich for correcting the initial confusion on the anyon statistics. M. O. is supported in part by Grant-in-Aid for Scientific Research, and a 21st Century COE Program at Tokyo Institute of Technology “Nanometer-Scale Quantum Physics,” both from MEXT of Japan. T. S. acknowledges support from NSF Grant No. DMR-0308945, funding from the NEC Corporation, the Alfred P. Sloan Foundation, and The Research Corporation.

-
- [1] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
 - [2] A. Yu. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003).
 - [3] N. Read and S. Sachdev, Phys. Rev. Lett. **66**, 1773 (1991); X. G. Wen, Phys. Rev. B **44**, 2664 (1991).
 - [4] T. Senthil and M. P. A. Fisher, Phys. Rev. B **62**, 7850 (2000); **63**, 134521 (2001).
 - [5] R. Moessner and S. L. Sondhi, Phys. Rev. Lett. **86**, 1881 (2001).
 - [6] G. Misguich, D. Serban, and V. Pasquier, Phys. Rev. Lett. **89**, 137202 (2002).
 - [7] L. Balents, M. P. A. Fisher, and S. M. Girvin, Phys. Rev. B **65**, 224412 (2002).
 - [8] O. I. Motrunich and T. Senthil, Phys. Rev. Lett. **89**, 277004 (2002).
 - [9] O. I. Motrunich, Phys. Rev. B **67**, 115108 (2003).
 - [10] M. Freedman, C. Nayak, K. Shtengel, K. Walker, and Z. Wang, Ann. Phys. (N.Y.) **310**, 428 (2004).
 - [11] X.-G. Wen, Int. J. Mod. Phys. B **5**, 1641 (1991).
 - [12] Y.-S. Wu, Y. Hatsugai, and M. Kohmoto, Phys. Rev. Lett. **66**, 659 (1991).
 - [13] R. Tao and Yong-Shi Wu, Phys. Rev. B **30**, R1097 (1984).
 - [14] X.-G. Wen and Q. Niu, Phys. Rev. B **41**, 9377 (1990).
 - [15] X.-G. Wen, E. Dagotto, and E. Fradkin, Phys. Rev. B **42**, 6110 (1990).
 - [16] C. Chamon, Phys. Rev. Lett. **94**, 040402 (2005).
 - [17] M. Oshikawa, Phys. Rev. Lett. **90**, 236401 (2003); **91**, 109901(E) (2003).
 - [18] A. Paramekanti and A. Vishwanath, Phys. Rev. B **70**, 245118 (2004).
 - [19] B. I. Halperin, Phys. Rev. Lett. **52**, 1583 (1984).
 - [20] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
 - [21] R. B. Laughlin, Phys. Rev. B **23**, R5632 (1981).
 - [22] M. Oshikawa, Phys. Rev. Lett. **84**, 1535 (2000).
 - [23] M. Oshikawa, Phys. Rev. Lett. **84**, 3370 (2000); T. Senthil, S. Sachdev, and M. Vojta, Phys. Rev. Lett. **90**, 216403 (2003).
 - [24] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004); Europhys. Lett. **70**, 824 (2005).