Oriented Closed Polyhedral Maps and the Kitaev Model

Kornél SZLACHÁNYI

Wigner Research Centre for Physics, Budapest, Hungary
E-mail: *szlachanyi.kornel@wigner.hu*

Abstract. A kind of combinatorial map, called arrow presentation, is proposed to encode the data of the oriented closed polyhedral complexes Σ on which the Hopf algebraic Kitaev model lives. We develop a theory of arrow presentations which underlines the role of the dual of the double $\mathcal{D}(\Sigma)^*$ of Σ as being the Schreier coset graph of the arrow presentation, explains the ribbon structure behind curves on $\mathcal{D}(\Sigma)^*$ and facilitates computation of holonomy with values in the algebra of the Kitaev model. In this way, we can prove ribbon operator identities for arbitrary f.d. C*-Hopf algebras and arbitrary oriented closed polyhedral maps. By means of a combinatorial notion of homotopy designed specially for ribbon curves, we can rigorously formulate "topological invariance" of states created by ribbon operators.

Key words: Hopf algebra; polyhedral map; quantum double; ribbon operator; topological invariance

2020 Mathematics Subject Classification: 05E99; 16T05; 81T25

1 Introduction

The toric code and its generalizations [1, 3, 12] belong to the class of models called topological orders which are thought to describe physical systems in which fault tolerant quantum computations can be realized. In its most general form, the model is a quantum system defined on the surface Σ of an abstract polyhedron having no particular symmetry and with quantum spins sitting on the edges of Σ and taking values in a finite-dimensional C*-Hopf algebra H. The mostly studied case is when H is the group algebra of a finite abelian group and Σ is the square lattice. This abelian Kitaev model exhibits charges classified by the quantum double $\mathcal{D}(H)$ that are created by semiinfinite string operators with the strings having no contribution to the energy, like Dirac strings. These charges obey anyon statistics [12, 18]. Much less is known beyond the abelian case. For general Hopf algebras, even the gauge theory interpretation is a challenge [15, 16]. In this paper, we define the general model starting from scratch. As for the lattice, we do not want to go beyond ribbon graphs used by [15] but wish to replace them by something truly 2-dimensional on which taking duals and doubles are straightforward operations. As a matter of fact, the dual of the double of Σ is as important as the dual of the double of H in the holonomy theory of [15]. Indeed, the dual of the double of the surface complex is but the "fattening" of the ribbon graph. But it is more than that. It is a sort of Cayley graph for the arrow presentation.

The idea of arrow presentation is well-known in the literature on graph embeddings, dessins d'enfants and permutation representations of triangle groups under various names. For embeddings of finite, connected, simple graphs these are the "rotation schemes" of [21, Section 6-6] and for finite, connected graphs these are the "combinatorial maps" of [13, Definition 1.3.23]. For more general maps see [11, 13]. A combinatorial map consists of a finite set A, the elements of which can be called directed edges, half edges, edge ends, darts or arrows, and of 3 permutations T_0, T_1, T_2 of A such that (1) $T_2 T_1 T_0 = \mathrm{id}_A$, (2) the group generated by the T_i-s acts

transitively on A (connectedness) and (3) the T_1 is an involution without fixed points. Such combinatorial data encode a "map", i.e., a kind of compact oriented 2-dimensional cellular complex Σ without boundary, in such a way that the i-cells of Σ correspond bijectively to the cycles of the permutation T_i. The arrow presentation we propose in this paper is a slight extension but also a specialization of combinatorial maps. We relax finiteness of A and connectedness but otherwise postulate stronger axioms on the permutations T_i. These stronger axioms lead to a map Σ in which there are no loop edges, no degree 1 vertices and no repetitions of any cell along the boundary of a face and along the coboundary of a vertex. Such a map Σ will be called an oriented closed polyhedral map (OCPM). Although the cardinality of A is not restricted, the connected OCPMs are always countable. By allowing infinite A-s such important examples as the square lattice on the plane and many other periodic and non-periodic tilings of the plane are given place among the connected OCPMs.

In Sections 2, 4 and 5, we sketch a theory of arrow presentations and of OCPMs. The emphasis is on the structure of the dual of the double, $\mathcal{D}(\Sigma)^*$, and on its ribbon curves. In Section 3, we define the algebra $\mathcal{M} = \mathcal{M}(\Sigma, H)$ of the Kitaev model on an OCPM Σ and for a finite-dimensional semisimple Hopf algebra H. In contrast to its representations the algebra itself can be presented without ever introducing an orientation of the edges. In Section 6, we introduce an $\mathcal{M}$-valued holonomy on curves of $\mathcal{D}(\Sigma)^*$ which is motivated by previous works [3, 15] although not depending on any relation to the Hopf algebra gauge theory of [16]. Here we find a notion of central deformation of curves which will be useful in later Sections. In Section 7, we launch the hard work of computing algebraic relations of ribbon operators, i.e., of the holonomies of ribbons. The results agree with the earlier results of [1, 5] obtained for group algebras or square lattices. More efficient relations exist not for the ribbon operators themselves but for their actions on vacuum states which is the topic of Sections 9 and 10. In order to prepare for them we have to introduce certain combinatorial notions of homotopy both for curves and ribbons. This is the content of Section 8. In Theorem 9.2, we are at last able to formulate topological invariance, or rather homotopy invariance, of states created by ribbon operators from the vacuum. Finding evidence for string localization of the charges in Section 10 the paper ends abruptly leaving open the question of existence of these charges, i.e., superselection sectors of the model.

2 Polyhedral maps

The main tool of this paper is the arrow presentation of certain surface complexes but arrow presentations themselves form an interesting mathematical structure.

Definition 2.1. An arrow presentation is a triple $\langle \mathrm{Arr}, T_0, T_2 \rangle$, where Arr is a set and for $i = 0, 2$ the $T_i \colon \mathrm{Arr} \to \mathrm{Arr}$ are bijective functions satisfying the following axioms:

(AP-1) The orbits $\mathcal{O}_i(a) := \left\{ T_i^n a \mid n \in \mathbb{N} \right\}$ are finite sets containing at least 2 elements for all $a \in \mathrm{Arr}$ and for $i = 0, 2$.

(AP-2) Introducing $T_1 := T_0 T_2$ we require that $T_1 T_1 = \mathrm{id}_{\mathrm{Arr}}$.

(AP-3) The intersection of a T_0-orbit and a T_2-orbit is either empty or a singleton.

The geometric meaning of the above definition will be unveiled gradually. Here is a rough picture of what is going on: The elements of Arr will be called arrows and will become the directed edges of a surface complex. Assuming that the complex is closed and oriented the meaning of the T_i transformations is this. Every directed edge a lies on the boundary of exactly 2 faces but only one of these faces, call it f, has compatible orientation with the direction of a. The orientation of f determines a cyclic order of its boundary edges and we define $T_2 a$ to be the successor of a within this cyclic order. Similarly one can define $T_0 a$ as the successor of a with

respect to rotating around the source vertex of a in counterclockwise direction. It is now easy to imagine what the composite operation T_1 does; it reverses the direction of the arrows.
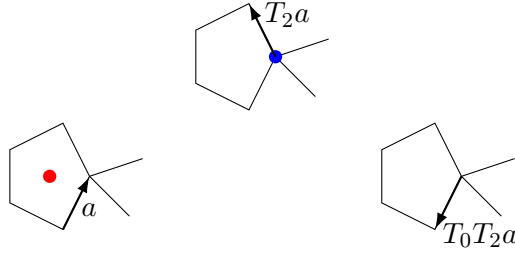
If we added the axiom $T_2^3 = \mathrm{id}_{\mathrm{Arr}}$, then we would be dealing with simplicial complexes but not doing so has the benefit of having a self-dual notion of complex, as we shall see later.

Now we derive some formal consequences of Definition 2.1. Note that the set Arr is not assumed to be finite.

Lemma 2.2. *Let* $\langle \mathrm{Arr}, T_0, T_2 \rangle$ *be an arrow presentation and let* $T_1 = T_0 T_2$. *Then*

(i) $T_2^{-1} = T_1 T_0$,

(ii) $T_0^{-1} = T_2 T_1$,

(iii) $\mathcal{O}_i(a) \cap \mathcal{O}_j(a) = \{a\}$ *for all* $a \in \mathrm{Arr}$ *and for* $0 \le i < j \le 2$,

(iv) $T_1(a) \ne a$ *for all* $a \in \mathrm{Arr}$.

Proof. (i) $T_1 T_0 T_2 = T_1 T_1 = \mathrm{id}_{\mathrm{Arr}}$ shows that T_2 has left inverse $T_1 T_0$. Since T_2 is invertible, $T_1 T_0$ is the inverse of T_2.

(ii) Using (i), we have $T_2 T_1 T_0 = \mathrm{id}_{\mathrm{Arr}}$. Therefore, $T_2 T_1$ is a left inverse of T_0. Since T_0 is invertible, $T_2 T_1$ is the inverse of T_0.

(iii) The $i = 0$, $j = 2$ case is trivial since a belongs to both orbits and by axiom (AP-3) no other element does so. In the $i = 0$, $j = 1$ case, assume $b \in \mathcal{O}_0(a) \cap \mathcal{O}_1(a)$. Then either $b = a$ or $b = T_1 a = T_0^m a$ for some m. In the latter case, $T_0^{-1} b = T_0^{m-1} a = T_2 a$ belongs to $\mathcal{O}_0(a) \cap \mathcal{O}_2(a) = \{a\}$ hence $T_2 a = a$ which contradicts (AP-1). Similarly, in the $i = 1$, $j = 2$ case a $c \in \mathcal{O}_1(a) \cap \mathcal{O}_2(a)$ which is not a must satisfy $c = T_1 a = T_2^n a$ for some n and therefore $c = T_0 T_2 a = T_2^{n-1} T_2 a$ belongs to $\mathcal{O}_0(T_2 a) \cap \mathcal{O}_2(T_2 a)$ hence $c = T_2 a$. Now we have $T_1 a = T_2 a$ which means that $T_2 a$ is a fixed point of T_0 which contradicts (AP-1).

(iv) Let $T_1 a = a$. Then $T_2 a = T_2 T_1 a = T_0^{-1} a$, so $T_2 a \in \mathcal{O}_0(a) \cap \mathcal{O}_2(a)$. By (iii), $T_2 a = a$ is a fixed point of T_2 which contradicts (AP-1). ∎

Out of an arrow presentation, we would like to build a 2-dimensional cellular complex, a kind of abstract polyhedron. The necessary data are the following. Let

$$\Sigma^i := \{O \subseteq \mathrm{Arr} \mid \exists a \in \mathrm{Arr}, \ \mathcal{O}_i(a) = O\} \tag{2.1}$$

be the set of T_i-orbits for $i = 0, 1, 2$. We call the elements $v \in \Sigma^0$ vertices, the elements $e \in \Sigma^1$ (undirected) edges and the elements $f \in \Sigma^2$ faces. For an edge $e = \{a, T_1 a\}$, we say that $u = \mathcal{O}_0(a)$ and $v = \mathcal{O}_0(T_1 a)$ are the endpoints, or boundary vertices, of e. In this case, we write $\mathrm{Bd}(e) = \{u, v\}$ and call it the boundary of e. The coboundary of e is defined by $\mathrm{Cb}(e) := \{\mathcal{O}_2(a), \mathcal{O}_2(T_1 a)\}$. More generally, we define

$$\mathrm{Bd}(x) := \{\mathcal{O}_{i-1}(a) \mid a \in x\} \qquad \forall x \in \Sigma^i, \quad i = 1, 2, \tag{2.2}$$

$$\mathrm{Cb}(x) := \{\mathcal{O}_{i+1}(a) \mid a \in x\} \qquad \forall x \in \Sigma^i, \quad i = 0, 1, \tag{2.3}$$

and we declare the boundaries of vertices and the coboundaries of faces to be empty. Denoting by Σ the disjoint union $\Sigma^0 \sqcup \Sigma^1 \sqcup \Sigma^2$ the Σ^i-s can be recovered by setting up a dimension function

$\dim \colon \Sigma \to \{0, 1, 2\}$. Then Bd and Cb become functions $\Sigma \to \mathbf{Fin}(\Sigma)$ to the set of finite subsets of Σ.

For a while, we forget about arrow presentations because first we have to formulate the requirements for a bunch of data, such as the $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ above, which qualifies it as a polyhedral complex.

Definition 2.3. The data $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ consisting of a set Σ and functions $\dim \colon \Sigma \to \{0, 1, 2\}$, $\mathrm{Bd}, \mathrm{Cb} \colon \Sigma \to \mathbf{Fin}(\Sigma)$ is called a closed polyhedral map (CPM) if the following axioms hold:

(CPM-1) For $s, t \in \Sigma$, if $s \in \mathrm{Bd}(t)$, then $\dim t = 1 + \dim s$.

(CPM-2) $s \in \mathrm{Bd}(t)$ if and only if $t \in \mathrm{Cb}(s)$ for all $s, t \in \Sigma$.

(CPM-3) $|\mathrm{Bd}(e)| = 2$ and $|\mathrm{Cb}(e)| = 2$ for all edges $e \in \Sigma^1$.

(CPM-4) For all $v \in \Sigma^0$, the elements of the vertex neighbourhood $\mathrm{Nb}(v) := \mathrm{Cb}(v) \cup \bigcup \{\mathrm{Cb}(e) \mid e \in \mathrm{Cb}(v)\}$ can be enumerated in such a way that it consists of edges $e_0, \ldots, e_{n-1}$ and faces $f_0, \ldots, f_{n-1}$ for some $n \geq 2$ with $\mathrm{Cb}(e_k) = \{f_k, f_{k+1}\}$ for $0 \leq k < n$, where $f_n = f_0$.

(CPM-5) For all $f \in \Sigma^2$, the elements of the face neighbourhood $\mathrm{Nb}(f) := \mathrm{Bd}(f) \cup \bigcup \{\mathrm{Bd}(e) \mid e \in \mathrm{Bd}(f)\}$ can be enumerated in such a way that it consists of edges $e_0, \ldots, e_{n-1}$ and vertices $v_0, \ldots, v_{n-1}$ for some $n \geq 2$ with $\mathrm{Bd}(e_k) = \{v_k, v_{k+1}\}$ for $0 \leq k < n$, where $v_n = v_0$.

For comparison of CPM-s with the complexes one obtains from combinatorial maps see Example 2.8.

The geometric realization of a CPM $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ is a 2-manifold $[\Sigma]$ without boundary that can be constructed from polygonal disks $[f]$ for each face $f \in \Sigma^2$ by sewing edges together. Axiom (CPM-5) ensures that the neighbourhood $\mathrm{Nb}(f)$ of a face f consists of n_f edges and n_f vertices for some $n_f \geq 2$ which form a circular chain so we can associate to the abstract cell f a topological closed disk $[f]$ the boundary of which is an n_f-gon and the edges and vertices of this polygon are labelled by the cells of $\mathrm{Nb}(f)$. Doing this for all $f \in \Sigma^2$ every edge label $e \in \Sigma^1$ occurs twice sitting on the boundary of two different polygonal disks, by axiom CPM-3. There is only one way to identify two edges of the same label which is compatible with the labels on their boundary vertices also by axiom CPM-3. Having done these identifications for all edges, we get a topological space $[\Sigma]$ in which every point has a neighbourhood homeomorphic to an open disk, hence a topological closed surface. The points where this is non-trivial are the vertices but for them axiom CPM-4 ensures open disk neighbourhoods.

$[\Sigma]$ is compact if and only if Σ is finite. The non-trivial 'only if' part follows from the existence of a cover of $[\Sigma]$ by open disks, with one disk for each cell, such that no proper subset of the cover is a cover.

Although Definition 2.3 deals only with closed complexes, i.e., ones without boundary and without coboundary, it is intuitively clear how to modify the definition in order to get polyhedral map (PM) satisfying the following theorem: *Every PM can be obtained from a CPM by discarding a set of faces (holes) and a set of vertices (punctures).*

Lemma 2.4. *For a vertex v and a face f of a CPM Σ the following conditions are equivalent:*

  (i) $v \in \mathrm{Nb}(f)$,

  (ii) $f \in \mathrm{Nb}(v)$,

  (iii) $\mathrm{Cb}(v) \cap \mathrm{Bd}(f) \neq \varnothing$.

Provided these conditions hold the intersection $\mathrm{Cb}(v) \cap \mathrm{Bd}(f)$ consists of precisely 2 edges.

Proof. The relation $v \in \mathrm{Nb}(f)$ is the square of the boundary relation, i.e., $\exists e \, (v \in \mathrm{Bd}(e) \wedge e \in \mathrm{Bd}(f))$. By (CPM-2), the transpose of Bd is Cb so (i) is equivalent to both (ii) and (iii). If $v \in \mathrm{Nb}(f)$, then axiom (CPM-5) shows explicitly the 2 coboundary edges of v that lie on the boundary of f. ■

If the CPM Σ arises from an arrow presentation as in equations (2.1) and (2.2), then the conditions of Lemma 2.4 can be supplemented by a fourth one, namely,

$$v \cap f \neq \varnothing \quad \Leftrightarrow \quad \mathrm{Cb}(v) \cap \mathrm{Bd}(f) \neq \varnothing. \tag{2.4}$$

In this case, the two edges can be explicitly given by

$$\mathrm{Cb}(v) \cap \mathrm{Bd}(f) = \{\mathcal{O}_1(a), \mathcal{O}_1(T_0 a)\} \qquad \text{where} \quad \{a\} = v \cap f. \tag{2.5}$$

Definition 2.5. A pair $\langle v, f \rangle \in \Sigma^0 \times \Sigma^2$ satisfying the conditions of Lemma 2.4 is called a site. Two sites $s_i = \langle v_i, f_i \rangle$, $i = 1, 2$, can be equal (if $v_1 = v_2$ and $f_1 = f_2$) or neighbour (if either $v_1 = v_2$ or $f_1 = f_2$ but not both) or disjoint (if none of the 2 equations holds).

The following equivalent conditions define connectedness of a CPM Σ:

- For every $u, v \in \Sigma^0$, there exists a sequence $(v_0, e_1, v_1, \ldots, e_n, v_n)$ of edges and vertices such that $v_0 = u$, $v_n = v$ and $\mathrm{Bd}(e_k) = \{v_{k-1}, v_k\}$ for $0 < k \leq n$.

- For every $p, q \in \Sigma^2$, there exists a sequence $(f_0, e_1, f_1, \ldots, e_n, f_n)$ of edges and faces such that $f_0 = p$, $f_n = q$ and $\mathrm{Cb}(e_k) = \{f_{k-1}, f_k\}$ for $0 < k \leq n$.

- For every pair s, t of sites, there exists a sequence $(s_i)_{i=0}^{n}$ of sites such that $s_0 = s$, $s_n = t$ and the $\{s_{i-1}, s_i\}$ are neighbours for $0 < i \leq n$.

If Σ is connected, then it is countable. As a matter of fact, such a Σ can be written as a countable union of finite subsets.

Axioms (CPM-4) and (CPM-5) offer an easy way to define orientation of CPMs. Axiom (CPM-5) allows to choose a cyclic order $[v_0, e_0, v_1, e_1, \ldots, v_{n-1}, e_{n-1}]$ on every face neighbourhood $\mathrm{Nb}(f)$ where $[x_1, \ldots, x_k]$ denotes sequences $(x_1, \ldots, x_k)$ modulo cyclic permutations. Since $n \geq 2$ by assumption, this cyclically ordered set has at least 4 elements. Therefore, there exists exactly two cyclic orders on $\mathrm{Nb}(f)$ satisfying (CPM-5). (If we did not include the vertices into $\mathrm{Nb}(f)$, we would be in trouble with the 2-gons.) Similarly, (CPM-4) allows to choose among the two possible cyclic orders on every vertex neighbourhood $\mathrm{Nb}(v)$.

Definition 2.6. Let $\langle v, f \rangle$ be a site. Cyclic orders $[f_0, e_0, f_1, e_1, \ldots, f_{m-1}, e_{m-1}]$ of $\mathrm{Nb}(v)$ and $[v_0, e_0, v_1, e_1, \ldots, v_{n-1}, e_{n-1}]$ of $\mathrm{Nb}(f)$ are called compatible if the predecessor and successor edges $e_p < f < e_s$ of f in $\mathrm{Nb}(v)$ and the predecessor and successor edges $e'_p < v < e'_s$ of v in $\mathrm{Nb}(f)$ are related by $e'_p = e_s$ and $e'_s = e_p$. Otherwise (i.e., if $e'_p = e_p$ and $e'_s = e_s$), they are called incompatible.

A system of cyclic orders given on all vertex and face neighbourhoods of Σ is called an orientation if for every site $\langle v, f \rangle$ on Σ the cyclic orders of $\mathrm{Nb}(v)$ and $\mathrm{Nb}(f)$ are compatible. A CPM Σ is called orientable if it has an orientation.

Theorem 2.7. *Given an arrow presentation $\langle \mathrm{Arr}, T_0, T_2 \rangle$ the associated data $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ defined by (2.1), (2.2) and (2.3) is a closed polyhedral map in the sense of Definition 2.3. Furthermore, the actions of T_0 and T_2 define canonical cyclic orders*

$$\begin{aligned}
[f_0, e_0, &\ldots, f_{m-1}, e_{m-1}] \\
&:= \left[\mathcal{O}_2(a), \mathcal{O}_1(T_0 a), \mathcal{O}_2(T_0 a), \mathcal{O}_1(T_0^2 a), \ldots, \mathcal{O}_2(T_0^{m-1} a), \mathcal{O}_1(a)\right]
\end{aligned} \tag{2.6}$$

on vertex neighbourhoods $\mathrm{Nb}(\mathcal{O}_0(a))$ *and*

$$[v_0, e_0, \ldots, v_{n-1}, e_{n-1}]$$
$$:= \left[\mathcal{O}_0(a), \mathcal{O}_1(a), \mathcal{O}_0(T_2 a), \mathcal{O}_1(T_2 a), \ldots, \mathcal{O}_0(T_2^{n-1} a), \mathcal{O}_1(T_2^{n-1} a)\right] \tag{2.7}$$

on face neighbourhoods $\mathrm{Nb}(\mathcal{O}_2(a))$. *These cyclic orders are independent of the choice of* a *within* $v = \mathcal{O}_0(a)$ *and within* $f = \mathcal{O}_2(a)$, *respectively, and constitute an orientation of* Σ *in the sense of Definition* 2.6. *In this way every arrow presentation determines an oriented closed polyhedral map.*

Proof. (CPM-1) is clear. The relation $s \in \mathrm{Bd}(t)$ means that there is an $a \in t$ such that $s = \mathcal{O}_{i-1}(a)$ where $i = \dim t$. But this is just the symmetric relation $s \cap t \neq \varnothing$. Inspecting also $\mathrm{Cb}(s)$, we get

$$s \in \mathrm{Bd}(t) \quad \Leftrightarrow \quad s \cap t \neq \varnothing \quad \Leftrightarrow \quad t \in \mathrm{Cb}(s), \tag{2.8}$$

which proves (CPM-2).

Every edge e is a 2-element set $\{a, T_1 a\}$ by (AP-2) and by Lemma 2.2 (iv). Therefore, both the boundary and the coboundary of e can contain at most 2 elements. The case $\mathcal{O}_i(a) = \mathcal{O}_i(T_1 a)$ for $i = 0$ or $i = 2$, however, is forbidden because it implies that $a \neq T_1 a \in \mathcal{O}_1(a) \cap \mathcal{O}_i(a)$ contradicting Lemma 2.2 (iii). This proves (CPM-3).

Let $v \in \Sigma^0$. Then the transformation T_0 gives a natural cyclic order on the set v of arrows, hence also on the set $\mathrm{Cb}(v)$ of edges. Let an $a \in v$ be fixed and set $e_k := \mathcal{O}_1(T_0^k a)$ for $0 \le k < n = |v|$. The coboundary of e_k is the pair $\mathrm{Cb}(e_k) = \left\{\mathcal{O}_2(T_0^k a), \mathcal{O}_2(T_1 T_0^k a)\right\}$. Noticing that Lemma 2.2 (i) implies that $\mathcal{O}_2(T_1 T_0^k a) = \mathcal{O}_2(T_0^{k-1} a)$, we will denote $\mathcal{O}_2(T_0^{k-1} a)$ by f_k and then get the desired formula $\mathrm{Cb}(e_k) = \{f_k, f_{k+1}\}$ for all k. It remains to show that there are exactly n different e_k-s and n different f_k-s. If $e_k = e_l$, then $T_0^k a$ is equal to either $T_0^l a$ or $T_1 T_0^l a$. In the first case $k = l \mod n$ because n is the size of $\mathcal{O}_0(a)$. In the second case $T_1 T_0^l a$ belongs to both the T_1-orbit and the T_0-orbit of $b = T_0^l a$. So $T_1 T_0^l a = b$ by Lemma 2.2 (iii) and b is a fixed point of T_1 contradicting Lemma 2.2 (iv). Thus, all e_k-s are different. In particular, $n = |\mathrm{Cb}(v)|$. Now assume $f_k = f_l$ for some $0 \le k < l < n$. Then, there is a power T_2^m such that $T_2^m b = T_0^{l-k} b$ for $b = T_0^{k-1} a$. Again by Lemma 2.2 (iii), we obtain $T_0^{l-k} b = b$ hence $k = l \mod n$. So all f_k-s are different and the vertex neighbourhood $\mathrm{Nb}(v)$ has the required circular structure. This proves (CPM-4).

The proof of (CPM-5) goes exactly as that of (CPM-4). Namely $\langle A, T_2^{-1}, T_0^{-1} \rangle$ is also an arrow presentation the T_1-orbits of which coincide with the original because $T_2^{-1} T_0^{-1} = T_0 T_2$. The T_0-orbits and T_2-orbits in turn are interchanged so (CPM-5) for $\langle \mathrm{Arr}, T_0, T_2 \rangle$ is equivalent to (CPM-4) for $\langle A, T_2^{-1}, T_0^{-1} \rangle$.

It remains to show compatibility of the cyclic orders given in (2.6) and (2.7) whenever v and f form a site. By (2.4) and Lemma 2.4, this means that $v \cap f \neq \varnothing$ so we can choose the a in both (2.6) and (2.7) to be the unique element of $v \cap f$. Then the 2 common edges of $\mathrm{Nb}(v)$ and $\mathrm{Nb}(f)$ are $\mathcal{O}_1(a)$ and $\mathcal{O}_1(T_0 a) = \mathcal{O}_1(T_2^{-1} a)$ the first of which is the predecessor of $\mathcal{O}_2(a) = f$ in $\mathrm{Nb}(v)$ and the successor of $\mathcal{O}_0(a) = v$ in $\mathrm{Nb}(f)$ while the second is the successor of f in $\mathrm{Nb}(v)$ and the predecessor of v in $\mathrm{Nb}(f)$. Therefore, they are compatible. ∎

A CPM Σ together with an orientation will be called an *oriented closed polyhedral map* (OCPM).

We will not formalize the converse of the above theorem which claims that every OCPM determines an arrow presentation; the idea has already been sketched below Definition 2.1. The two constructions AP $\longrightarrow$ OCPM and OCPM $\longrightarrow$ AP establish a bijective correspondence between (isomorphism classes of) oriented closed polyhedral maps and arrow presentations. The discussion of isomorphisms is postponed until Section 4.

Example 2.8. In order to illustrate the difference between combinatorial maps and arrow presentations, in particular, the role of axiom (AP-3), let us consider three maps on the torus:

$$\Sigma_{1\times 1} = \qquad \begin{array}{l} A = \{1, 2, \bar{1}, \bar{2}\}, \\ T_0 = (12\bar{1}\bar{2}), \\ T_2 = (12\bar{1}\bar{2}), \end{array}$$

$$\Sigma_{1\#2} = \qquad \begin{array}{l} A = \{1, 2, 3, 4, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}, \\ T_0 = (12\bar{3}\bar{4})(\bar{1}234), \\ T_2 = (14\bar{3}\bar{2})(32\bar{1}\bar{4}), \end{array}$$

$$\Sigma_{2\times 2} = \qquad \begin{array}{l} A = \{1, 2, 3, 4, 5, 6, 7, 8, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}, \bar{7}, \bar{8}\}, \\ T_0 = (12\bar{3}\bar{6})(34\bar{1}\bar{8})(56\bar{7}\bar{2})(78\bar{5}\bar{4}), \\ T_2 = (14\bar{5}\bar{2})(32\bar{7}\bar{4})(58\bar{1}\bar{6})(76\bar{3}\bar{8}). \end{array}$$

The arrows are named $1, \ldots, n, \bar{1}, \ldots, \bar{n}$ with $\bar{k}$ being the reverse of k. The permutations T_i are given by their cycle decompositions. So, in each of the cases $T_1 = (1\bar{1}) \ldots (n\bar{n})$.

$\Sigma_{1\times 1}$ arises from embedding the figure 8 graph (of 1 vertex and 2 loops) into the torus so that the two loops become a pair of meridian and longitude. $\Sigma_{1\#2}$ arises from embedding ⬤⟨⟩⬤ into the torus in the following way. First, we draw a length 2 longitude and two disjoint length 1 meridians of distance 1 from each other. Then, we cut along the longitude, shift it by 1 and glue it back (half Dehn twist). $\Sigma_{2\times 2}$ is also an embedding of a graph which can be best drawn on the torus itself.

It is easy to see that in each of the three cases the data $\langle A, T_0, T_1, T_2 \rangle$ obey the three axioms of combinatorial maps mentioned in the introduction. Moreover, neither of the T_i have a fixed point so also the axioms (AP-1) and (AP-2) are fulfilled. In the map $\Sigma_{1\times 1}$ we see the same face on both hand sides of an edge (isthmus) because there is only one face. Also the edges are bordered by one and the same vertex. So this example is breaking the axioms (CPM-3), (CPM-4) and (CPM-5). In terms of the permutations, we see the weird identity $T_0 = T_2$ implying immediately that (AP-3) fails. In the map $\Sigma_{1\#2}$, there are no loops and no isthmuses but there are repeated faces in the vertex neighbourhoods and repeated vertices in the face neighbourhoods. So (CPM-4) and (CPM-5) are broken. Both the first cycle of T_0 and the second cycle of T_2 contains 2 and $\bar{4}$ which means that (AP-3) fails. This is not an arrow presentation, either. The map $\Sigma_{2\times 2}$ is the only one among the three which is an OCPM and the only one for which $\langle A, T_0, T_2 \rangle$ is an arrow presentation. This example proves independence of the axiom (AP-3) and shows also that in order to produce a map with given genus arrow presentations need more arrows than combinatorial maps do.

Let us mention briefly two important constructions of CPMs which work also for non-orientable Σ. See Figure 1.

Definition 2.9. Let $\Sigma = \langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ be a closed polyhedral map. Its dual is

$$\Sigma^* := \langle \Sigma, 2 - \dim, \mathrm{Cb}, \mathrm{Bd} \rangle \tag{2.9}$$

and its double $\mathcal{D} = \mathcal{D}(\Sigma) = \langle \mathcal{D}^0, \mathcal{D}^1, \mathcal{D}^2, \mathbf{Bd}, \mathbf{Cb} \rangle$ is defined by

$$\mathcal{D}^0 := \Sigma^0 \sqcup \Sigma^1 \sqcup \Sigma^2,$$
$$\mathcal{D}^1 := \{ \langle v, e \rangle \in \Sigma^0 \times \Sigma^1 \mid v \in \mathrm{Bd}(e) \} \sqcup \{ \langle e, f \rangle \in \Sigma^1 \times \Sigma^2 \mid e \in \mathrm{Bd}(f) \},$$
$$\mathcal{D}^2 := \{ \langle v, f \rangle \in \Sigma^0 \times \Sigma^2 \mid \mathrm{Cb}(v) \cap \mathrm{Bd}(f) \neq \varnothing \},$$
$$\mathbf{Bd}(\langle s, t \rangle) := \{s, t\} \qquad \text{for } \langle s, t \rangle \in \mathcal{D}^1,$$
$$\mathbf{Bd}(\langle v, f \rangle) := \{ \langle v, e_1 \rangle, \langle v, e_2 \rangle, \langle e_1, f \rangle, \langle e_2, f \rangle \},$$

where $\{e_1, e_2\} = \mathrm{Cb}(v) \cap \mathrm{Bd}(f)$ for $\langle v, f \rangle \in \mathcal{D}^2$.
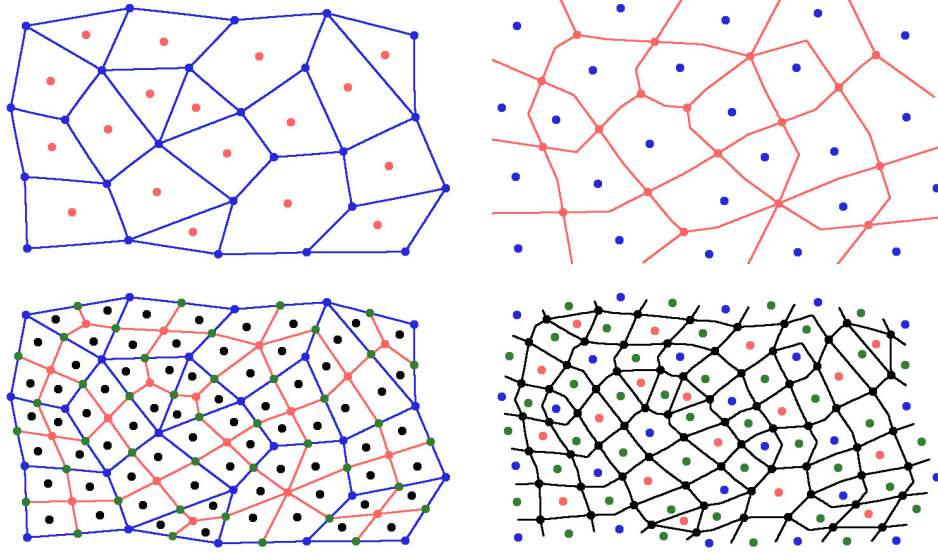
Figure 1. A planar region of a CPM Σ and the same region of Σ^*, $\mathcal{D}(\Sigma)$ and $\mathcal{D}(\Sigma)^*$. Blue dots are the vertices, red dots are the faces. In the 2nd row the green dots are the edges, black ones are the sites.

The coboundary **Cb** is then obtained by transposing the relation $x \in \mathbf{Bd}(y)$.

Notice that the edges of $\mathcal{D}(\Sigma)$ are naturally divided into two classes: the ve-type edges and the ef-type edges. Pairs of ve edges organize into the edges of Σ and pairs of ef edges do so for the dual complex Σ^*, explaining the name "double" of Σ. The double $\mathcal{D}(\Sigma)$ is always free of multiple edges and loops.

Notice that every face of $\mathcal{D}(\Sigma)$ is a quadrangle. The faces of the double are precisely the sites of Definition 2.5. They are just the vertices of the dual of the double. Therefore, every vertex of $\mathcal{D}(\Sigma)^*$ is 4-valent.

Example 2.10. The minimal OCPM. Let $\Sigma_{\min}$ be the CPM consisting of 2 vertices, 2 edges and 2 faces. Topologically $\Sigma_{\min}$ is a sphere made up of two hemispheres f_1, f_2 separated by an equator consisting of two edges e_1, e_2 and two vertices v_1, v_2 such that $\mathrm{Bd}(e_1) = \mathrm{Bd}(e_2) = \{v_1, v_2\}$ and $\mathrm{Cb}(e_1) = \mathrm{Cb}(e_2) = \{f_1, f_2\}$. Fixing a to be one of the 4 arrows, we give an orientation to $\Sigma_{\min}$ by the arrow presentation

$$\langle \{a, T_1 a, T_0 a, T_2 a\}, T_0, T_2 \rangle,$$

where the action of T_0 and T_2 are fixed by the relations $T_0^2 = \mathrm{id}$, $T_2^2 = \mathrm{id}$. Since also $(T_0 T_2)^2 = \mathrm{id}$, this is the standard presentation of the left regular G-set over the Klein four-group $G = Z_2 \times Z_2$. It follows that if $\langle A, T_0, T_2 \rangle$ is an arrow presentation of a connected OCPM Σ such that there is an $a \in A$ satisfying $T_0^2 a = a = T_2^2 a$ then $\Sigma = \Sigma_{\min}$. The dual of the double
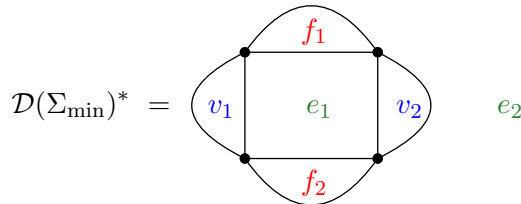
is another spherical OCPM the faces of which are the cells of $\Sigma_{\min}$.

In the rest of the paper, we shall work mainly with connected oriented closed polyhedral maps and freely mix the languages of $\langle \mathrm{Arr}, T_0, T_2 \rangle$ and $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$. We write $\Sigma :: P$ if $P = \langle \mathrm{Arr}, T_0, T_2 \rangle$ is a presentation of the OCPM Σ.

3 Kitaev's model over a Hopf algebra

Before launching the definition of the model, let us fix our conventions and notations about Hopf algebras. We work with finite-dimensional Hopf algebras $\langle H, \cdot, 1, \Delta, \varepsilon, S \rangle$ over a field K. Later in Section 10, K will be the complex field. Let H^* denote the dual Hopf algebra structure on the space of linear functions $H \to K$. The evaluation of $\varphi \in H^*$ on $h \in H$ is denoted either by $\varphi(h)$ or, if φ is a long expression, by $\langle \varphi, h \rangle$ or even by $\langle h, \varphi \rangle$ since H^{**} can be identified with H. We use the standard Hopf algebraist notation $\Delta(h) = h_{(1)} \otimes h_{(2)}$ for the coproduct of an element which is a concise notation for a linear combination $\sum_i h_{(1)i} \otimes h_{(2)_i}$. We often push this convention to the extreme and write, for example, $\varphi \otimes h \in H^* \otimes H$ but mean a general linear combination $\sum_i \varphi_i \otimes h_i$ of some $\varphi_i \in H^*$ and $h_i \in H$. In want of enough letters and due to abundance of *-s in this paper the antipode and counit of H^* will be denoted also by S and ϵ, respectively.

A basis $\{x_i\}$ of H and a basis $\{\xi_i\}$ of H^* are called dual bases if $\xi_i(x_j) = \delta_{i,j}$. For the indices of these bases, we use Einstein's summation convention: repeated indices are summed over. In this sense, we have the useful identities

$$x_i \otimes x_j \otimes \xi_i \xi_j = x_{k(1)} \otimes x_{k(2)} \otimes \xi_k, \qquad x_i x_j \otimes \xi_i \otimes \xi_j = x_k \otimes \xi_{k(1)} \otimes \xi_{k(2)},$$
$$S(x_i) \otimes \xi_i = x_i \otimes S(\xi_i),$$
$$\langle \varphi, x_i \rangle \xi_i = \varphi, \qquad\qquad\qquad x_i \langle \xi_i, y \rangle = y.$$

The Hopf algebra H is a left and right module algebra over H^* with respect to the actions $\varphi \rightharpoonup h := h_{(1)} \varphi(h_{(2)})$ and $h \leftharpoonup \varphi := \varphi(h_{(1)}) h_{(2)}$, respectively.

The Drinfeld double $\mathcal{D}(H)$, a.k.a. the quantum double, is the following Hopf algebra structure on the vector space $H^* \otimes H$: The multiplication is

$$(\varphi \otimes g)(\psi \otimes h) = \varphi \psi_{(2)} \otimes g_{(2)} h \cdot \psi_{(3)}(g_{(1)}) \psi_{(1)}(S^{-1}(g_{(3)}))$$

with unit $1_{\mathcal{D}} = \varepsilon \otimes 1$ and the comultiplication is

$$\Delta(\psi \otimes h) = (\psi_{(2)} \otimes h_{(1)}) \otimes (\psi_{(1)} \otimes h_{(2)})$$

with counit $\varepsilon_{\mathcal{D}}(\psi \otimes h) = \psi(1)\varepsilon(h)$. The antipode is given by the formula

$$S_{\mathcal{D}}(\varphi \otimes h) = S^{-1}(\varphi_{(2)}) \otimes S(h_{(2)}) \cdot \varphi_{(1)}(h_{(3)}) \varphi_{(3)}(S^{-1}(h_{(1)})).$$

Equipped with the R-matrix

$$R \equiv R_1 \otimes R_2 := (\varepsilon \otimes x_i) \otimes (\xi_i \otimes 1)$$

the double $\mathcal{D}(H)$ is a quasitriangular Hopf algebra for any finite-dimensional H.

The dual Hopf algebra of $\mathcal{D}(H)$ is denoted by $\mathcal{D}(H)^*$. It is the vector space $H \otimes H^*$ equipped with algebra structure

$$(h \otimes \varphi)(g \otimes \psi) = gh \otimes \varphi\psi, \qquad 1_{\mathcal{D}^*} = 1 \otimes \varepsilon, \tag{3.1}$$

with coalgebra structure

$$\Delta_{\mathcal{D}^*}(h \otimes \varphi) = (h_{(1)} \otimes \xi_i \varphi_{(1)} \xi_j) \otimes (S^{-1}(x_j) h_{(2)} x_i \otimes \varphi_{(2)}), \qquad \varepsilon_{\mathcal{D}^*}(h \otimes \varphi) = \varepsilon(h)\varphi(1) \tag{3.2}$$

and with antipode

$$S_{\mathcal{D}^*}(h \otimes \varphi) = x_j S^{-1}(h) S^{-1}(x_i) \otimes \xi_i S(\varphi) \xi_j = (x_i \otimes \xi_i)^{-1}(S^{-1}(h) \otimes S(\varphi))(x_j \otimes \xi_j). \tag{3.3}$$

An element $h \in H$ is called cocommutative if $h_{(2)} \otimes h_{(1)} = h_{(1)} \otimes h_{(2)}$. The set $\mathrm{Cocom}\, H$ of cocommutative elements is a subalgebra (but not a subcoalgebra, in general) of H.

If H is involutive, i.e., $S^2 = \mathrm{id}_H$, then so is H^*, $\mathcal{D}(H)$ and $\mathcal{D}(H)^*$. Involutive Hopf algebras are very close to being semisimple. By a theorem of Larson and Radford, finite-dimensional involutive Hopf algebras are semisimple provided the ground field K has characteristic 0 [14]. The Hopf algebra H is semisimple if and only if it contains an element i satisfying $hi = i\varepsilon(h)$, $\forall h \in H$ and $\varepsilon(i) = 1$. If it exists, such an element is unique and called the Haar integral. The Haar integral further satisfies $S(i) = i$ and $i_{(2)} \otimes i_{(1)} = i_{(1)} \otimes i_{(2)}$. So $i \in \mathrm{Cocom}\, H$. If H is semisimple, then so is H^*, $\mathcal{D}(H)$ and $\mathcal{D}(H)^*$ with respective Haar integrals ι, $i_{\mathcal{D}} = \iota \otimes i$ and $i_{\mathcal{D}^*} = i \otimes \iota$.

After this preparation, we can define the algebra of the model as follows. Let $\Sigma :: \langle \mathrm{Arr}, T_0, T_2 \rangle$ be an OCPM and assume that Arr is finite. Let H be a finite-dimensional Hopf algebra which is involutive. We construct the algebra $\mathcal{M} = \mathcal{M}(\Sigma, H)$ as the tensor product algebra

$$\mathcal{M}(\Sigma, H) = \bigotimes_{e \in \Sigma^1} \mathcal{M}_e,$$

where the algebras $\mathcal{M}_e$ can be presented by a redundant set of generators

$$P_a(h),\ Q_a(\varphi), \qquad \text{where} \quad h \in H,\ \varphi \in H^*,\ a \in e$$

and by the relations

$$P_a(h),\ Q_a(\varphi) \text{ are linear in } h \text{ and } \varphi, \tag{3.4}$$
$$P_a(h)P_a(k) = P_a(hk), \qquad P_a(1) = \mathbf{1}_{\mathcal{M}}, \tag{3.5}$$
$$Q_a(\varphi)Q_a(\psi) = Q_a(\varphi\psi), \qquad Q_a(\varepsilon) = \mathbf{1}_{\mathcal{M}}, \tag{3.6}$$
$$Q_a(\varphi)P_a(h) = P_a\big(\varphi_{(1)} \rightharpoonup h\big)Q_a\big(\varphi_{(2)}\big), \tag{3.7}$$

if $\{x_i\}$ and $\{\xi_i\}$ are dual bases, then

$$P_{T_1 a}(h) = P_a(x_j S(h) x_i) Q_a(S(\xi_i)\xi_j), \tag{3.8}$$
$$Q_{T_1 a}(\varphi) = P_a(x_j S(x_i)) Q_a(\xi_i S(\varphi)\xi_j) \tag{3.9}$$

for all $h, k \in H$, $\varphi, \psi \in H^*$ and $a \in e$. The reader may observe that P_a and Q_a are nothing but the triangle operators L_+ and T_+ of [12], although defined as abstract operators without representation on a (Hilbert) space. (The L_- and T_- are the L_+ and T_+ of the opposite arrow $T_1 a$, as explained below.)

The first 4 relations tell us that for each $a \in \mathrm{Arr}$ the algebra generated by $P_a(H)$ and $Q_a(H^*)$ is the Heisenberg double of H, also called the smash product algebra $H \# H^*$.

The redundancy of the generators is made explicit by the last 2 relations so what needs explanation is that why the $P_{T_1 a}$ and $Q_{T_1 a}$ obey the same relations as the P_a and Q_a. This is closely related to involutivity of the antipode. But for the sake of explanation let H be any finite-dimensional Hopf algebra. Recall [17] that the smash product algebra is isomorphic to the full endomorphism algebra $\mathrm{End}\, H$. In fact, there are two canonical isomorphisms

$$\mathcal{L}\colon\ H \otimes H^* \to \mathrm{End}\, H, \qquad \mathcal{L}(h \otimes \varphi)(x) = h(\varphi \rightharpoonup x), \tag{3.10}$$
$$\mathcal{R}\colon\ H \otimes H^* \to \mathrm{End}\, H, \qquad \mathcal{R}(h \otimes \varphi)(x) = \big(x \leftharpoonup S^{-1}(\varphi)\big)S(h) \tag{3.11}$$

related by $\mathcal{R}(\cdot) = S \circ \mathcal{L}(\cdot) \circ S^{-1}$. Introducing $p(h) := \mathcal{L}(h \otimes \varepsilon)$, $q(\varphi) := \mathcal{L}(1 \otimes \varphi)$ and $\bar{p}(h) : \mathcal{R}(h \otimes \varepsilon)$, $\bar{q}(\varphi) := \mathcal{R}(1 \otimes \varphi)$, we see that the operators p and q satisfy the same relations

as $\bar{p}$ and $\bar{q}$, namely (3.5), (3.6) and (3.7). It is clear that $\bar{p}$ and $\bar{q}$ can be expressed in terms of p and q and vice versa; we only have to compute

$$\mathcal{L}^{-1} \circ \mathcal{R}(h \otimes \varphi) = x_j S(h) S^{-1}(x_i) \otimes \xi_i S^{-1}(\varphi) \xi_j, \tag{3.12}$$

$$\mathcal{R}^{-1} \circ \mathcal{L}(h \otimes \varphi) = x_j S^{-1}(h) S^{-1}(x_i) \otimes \xi_i S(\varphi) \xi_j, \tag{3.13}$$

which are identical expressions precisely if S is involutive. Specializing them by substituting $\varphi = \varepsilon$ or $h = 1$, we obtain the required expressions (3.8) and (3.9). At the end we did not have to choose whether P_a and Q_a were represented by $\mathcal{L}$ and $P_{T_1 a}$ and $Q_{T_1 a}$ by $\mathcal{R}$ or vice versa. That is to say, with the given generators, the presentation of $\mathcal{M}$ does not require any orientation of the edges: a and $T_1 a$ are treated symmetrically as well as the P, Q operators associated with them. In representations, however, such a choice has to be made.

Summarizing, we have a full matrix algebra $\mathcal{M}_e \cong H \# H^* \cong \operatorname{End} H$ for each edge e of Σ and $\mathcal{M}(\Sigma)$ is the tensor product of these, so itself is a full matrix algebra. The long definition above should therefore be considered not just as the definition of the algebra $\mathcal{M}$ but as the definition of $\mathcal{M}$ together with a distinguished set of generators. This set of generators strongly influences what we shall ask and what we can answer in this model.

Some consequences of the defining relations of $\mathcal{M}$ are the following:

$$P_{T_1 a}(k) \quad \text{commutes with} \quad P_a(h), \tag{3.14}$$

$$Q_{T_1 a}(\psi) \quad \text{commutes with} \quad Q_a(\varphi), \tag{3.15}$$

$$P_a(h) Q_a(\varphi) = Q_a\big(\varphi \leftharpoonup S(h_{(2)})\big) P_a\big(h_{(1)}\big), \tag{3.16}$$

$$Q_a(\varphi) P_{T_1 a}(h) = P_{T_1 a}\big(h \leftharpoonup S(\varphi_{(2)})\big) Q_a\big(\varphi_{(1)}\big), \tag{3.17}$$

$$P_{T_1 a}(h) Q_a(\varphi) = Q_a\big(h_{(1)} \rightharpoonup \varphi\big) P_{T_1 a}\big(h_{(2)}\big) \tag{3.18}$$

for all $h, k \in H$, $\varphi, \psi \in H^*$ and $a \in \operatorname{Arr}$. (3.16) is the inverse of (3.7). (3.17) can be most easily obtained using the $\mathcal{L}$-$\mathcal{R}$ representations mentioned above. (3.18) is related to (3.17) as (3.16) does to (3.7). Fortunately, we will never need the complicated relations (3.8) and (3.9); the (3.14), (3.15) and (3.18) can be used as perfect substitutes for them.

The fact that the degrees of freedom are attached to the arrows of Σ suggests that we should treat the Kitaev model as a gauge theory. Since there are two fields P_a and Q_a at each arrow satisfying Heisenberg double relations means something like having a quantized gauge theory with an involutive finite-dimensional Hopf algebra H playing the role of the gauge group. The field Q_a can be interpreted as an exponential form of a vector potential and its conjugate momentum P_a as an exponentiated electric field. In spite of this interpretation, we are not going to follow the standard strategy of gauge theory of declaring the H-gauge invariant subalgebra as the algebra of observables. Following [12], we shall treat to whole $\mathcal{M}$ as consisting of physically realizable operators. There exist certain copies of the Drinfeld double $\mathcal{D}(H)$ within $\mathcal{M}$ as subalgebras. Their induced action on $\mathcal{M}$ will be called "gauge transformations" with the warning that we are just borrowing a word from gauge theory and applying it to a new situation in which their role is yet to be explored.

In fact, there are plenty of ways to embed the double into $\mathcal{M}$ as a subalgebra.

Lemma 3.1. *Let $\mathbf{a} = (a_1, \ldots, a_m)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ be sequences of arrows. Define*

$$G_{a_1, \ldots, a_m}(h) := P_{a_1}\big(h_{(1)}\big) \ldots P_{a_m}(h_{(m)}), \qquad h \in H, \tag{3.19}$$

$$F_{b_1, \ldots, b_n}(\psi) := Q_{b_1}\big(\psi_{(1)}\big) \ldots Q_{b_n}(\psi_{(n)}), \qquad \psi \in H^*. \tag{3.20}$$

Assume that

(i) *neither the sequence $\mathcal{O}_1(a_1), \ldots, \mathcal{O}_1(a_m)$ nor the sequence $\mathcal{O}_1(b_1), \ldots, \mathcal{O}_1(b_n)$ contains repeated edges,*

(ii) $a_m = b_1$, $b_n = T_1 a_1$,

(iii) *if $\mathcal{O}_1(a_i) = \mathcal{O}_1(b_j)$ then either $i = 1$ and $j = n$ or $i = m$ and $j = 1$.*

Then the map $\mathbf{D}(\psi \otimes h) := F_\mathbf{b}(\psi) G_\mathbf{a}(h)$ is an algebra homomorphism $\mathcal{D}(H) \to \mathcal{M}$. That is to say,

$$G_\mathbf{a}(g) G_\mathbf{a}(h) = G_\mathbf{a}(gh),$$
$$F_\mathbf{b}(\varphi) F_\mathbf{b}(\psi) = F_\mathbf{b}(\varphi\psi),$$
$$G_\mathbf{a}(h) F_\mathbf{b}(\psi) = F_\mathbf{b}(\psi_{(2)}) G_\mathbf{a}(h_{(2)}) \cdot \psi_{(3)}(h_{(1)}) \psi_{(1)}(S(h_{(3)}))$$

holds for all $g, h \in H$ and for all $\varphi, \psi \in H^*$. If H is semisimple, then $\mathbf{D}$ is injective.

Proof. The first two relations are simple consequences of multiplicativity of the coproduct and of assumption (i). The third can be proven by writing both F and G as a product of 3 terms (disregarding implicit linear combinations hidden in the coproducts), two side terms and one middle term where the middle term commutes with all other P-s and Q-s due to assumption (iii). The essence of the calculation is the exchange relations (3.16) and (3.18) applied to the side terms:

$$\begin{aligned}
G_\mathbf{a}(h) F_\mathbf{b}(\psi) &= P_{a_1}(h_{(1)}) P_{\ldots}(h_{(2)}) P_{a_m}(h_{(3)}) \cdot Q_{b_1}(\psi_{(1)}) Q_{\ldots}(\psi_{(2)}) Q_{b_n}(\psi_{(3)}) \\
&= P_{a_1}(h_{(1)}) P_{\ldots}(h_{(2)}) Q_{b_1}(\psi_{(1)} \leftharpoonup S(h_{(4)})) P_{a_m}(h_{(3)}) Q_{\ldots}(\psi_{(2)}) Q_{b_n}(\psi_{(3)}) \\
&= Q_{b_1}(\psi_{(1)} \leftharpoonup S(h_{(4)})) Q_{\ldots}(\psi_{(2)}) P_{a_1}(h_{(1)}) Q_{b_n}(\psi_{(3)}) P_{\ldots}(h_{(2)}) P_{a_m}(h_{(3)}) \\
&= Q_{b_1}(\psi_{(1)} \leftharpoonup S(h_{(5)})) Q_{\ldots}(\psi_{(2)}) Q_{b_n}(h_{(1)} \rightharpoonup \psi_{(3)}) P_{a_1}(h_{(2)}) P_{\ldots}(h_{(3)}) P_{a_m}(h_{(4)}) \\
&= \psi_{(1)}(S(h_{(3)})) \cdot F_\mathbf{b}(\psi_{(2)}) G_\mathbf{a}(h_{(2)}) \cdot \psi_{(3)}(h_{(1)}).
\end{aligned}$$

It remains to show injectivity of $\mathbf{D}$. Since the image of $\mathbf{D}$ lies in the tensor product of two copies of $H \# H^*$ at the edges $e_1 := \mathcal{O}_1(a_1)$, $e_2 := \mathcal{O}_1(a_m)$ and copies of H or H^* at the intermediate edges $\mathcal{O}_1(a_2), \ldots, \mathcal{O}_1(a_{m-1})$ and $\mathcal{O}_1(b_2), \ldots, \mathcal{O}_1(b_{n-1})$, respectively, if we compose $\mathbf{D}$ with the counit of H, resp. H^* at the intermediate places we obtain the $m = n = 2$ version of $\mathbf{D}$. So injectivity will follow from that of the $m = n = 2$ case. Further composing with representation $\mathcal{L}$ at both e_1 and e_2, the map becomes a representation on $H \otimes H$ given by

$$\mathbf{D}'(\psi \otimes h) \cdot (g \otimes k) = (h_{(1)} g) \leftharpoonup S(\psi_{(2)}) \otimes \psi_{(1)} \rightharpoonup (h_{(2)} k).$$

Let $i \in H$ be the Haar integral of H. We claim that $i \otimes 1$ is a cyclic vector of this representation. As a matter of fact, if $\mathbf{D}'(\psi \otimes h) \cdot (i \otimes 1) = i \leftharpoonup S(\psi_{(2)}) \otimes \psi_{(1)} \rightharpoonup h$ is zero then so is $\lambda \otimes v := S(\psi_{(2)}) \otimes \psi_{(1)} \rightharpoonup h$, and therefore

$$0 = S(\lambda_{(1)}) \otimes \lambda_{(2)} \rightharpoonup v = \psi_{(3)} \otimes S(\psi_{(2)}) \rightharpoonup (\psi_{(1)} \rightharpoonup h) = \psi \otimes h.$$

Hence, the kernel of $\mathbf{D}'$, containing the kernel of $\mathbf{D}$, is zero.   ∎

Note that the sequences of arrows on which (3.20) and (3.19) are defined are arbitrary sets of arrows with an ordering, they do not have to respect any neighbourhood relations on the surface complex Σ. Of course, $\mathcal{M}$ is just the tensor product of edge algebras, so knows nothing about how the edges organize into a complex.

By some vague analogy with gauge theory, we nevertheless consider those sequences $\mathbf{a}$, $\mathbf{b}$ which are closed curves on Σ: a Wilson loop of Q operators and a dual Wison loop of P operators. (Although it works for the minimal loops, taking larger ones we have to realize that

these are not going to define a holonomy!) The easiest way to construct such loops is to take the T_2-orbits or the T_0-orbits of a fixed arrow. For $a \in \mathrm{Arr}$, we define

$$G_a(h) := P_{T_0 a}(h_{(1)}) P_{T_0^2 a}(h_{(2)}) \cdots P_{T_0^m a}(h_{(m)}), \qquad m = |\mathcal{O}_0(a)|, \tag{3.21}$$

$$F_a(\varphi) := Q_a(\varphi_{(1)}) Q_{T_2 a}(\varphi_{(2)}) \cdots Q_{T_2^{n-1} a}(\varphi_{(n)}), \qquad n = |\mathcal{O}_2(a)|. \tag{3.22}$$

G_a is a sort of divergence of the electric field at the vertex $\mathcal{O}_0(a)$ so we call it the Gauss' law operator. F_a is sort of curl of the vector potential at the face $\mathcal{O}_2(a)$ so we call it the magnetic flux operator. By Lemma 3.1, the operators

$$\mathbf{D}_a(\varphi \otimes h) := F_a(\varphi) G_a(h), \qquad \varphi \otimes h \in \mathcal{D}(H)$$

provide homomorphic images of the double in $\mathcal{M}$ for each $a \in \mathrm{Arr}$, and therefore

$$(\psi \otimes k) \underset{a}{\triangleright} M := \mathbf{D}_a\big((\psi \otimes k)_{(1)}\big) M \, \mathbf{D}_a\big(S_{\mathcal{D}}\big((\psi \otimes k)_{(2)}\big)\big)$$
$$= F_a(\psi_{(2)}) G_a(k_{(1)}) \, M \, G_a\big(S(k_{(2)})\big) F_a\big(S(\psi_{(1)})\big),$$
$$M \underset{a}{\triangleleft} (\psi \otimes k) := \mathbf{D}_a\big(S_{\mathcal{D}}\big((\psi \otimes k)_{(1)}\big)\big) M \, \mathbf{D}_a\big((\psi \otimes k)_{(2)}\big)$$
$$= G_a\big(S(k_{(1)})\big) F_a\big(S(\psi_{(2)})\big) \, M \, F_a(\psi_{(1)}) G_a(k_{(2)})$$

define a left $\mathcal{D}(H)$-module algebra and a right $\mathcal{D}(H)$-module algebra structure on $\mathcal{M}$ for each $a \in \mathrm{Arr}$. For different arrows, however, these actions do not necessarily commute. The F_a loop operators are supported on the oriented boundary of the face $\mathcal{O}_2(a)$, so $F_a(\psi)$ and $F_b(\varphi)$ commute whenever $\mathcal{O}_2(a) \neq \mathcal{O}_2(b)$ because of (3.15). Similarly, the G_a-s are supported on the oriented coboundary (a star) of the vertex $\mathcal{O}_0(a)$, so $G_a(k)$ and $G_b(h)$ commute whenever $\mathcal{O}_0(a) \neq \mathcal{O}_0(b)$ because of (3.14). It follows that if $\mathcal{O}_0(a) \neq \mathcal{O}_0(b)$ and $\mathcal{O}_2(a) \neq \mathcal{O}_2(b)$ then the images of $\mathbf{D}_a$ and $\mathbf{D}_b$ commute and so do the actions $\triangleright_a$ or $\triangleleft_a$ with $\triangleright_b$ or $\triangleleft_b$. Thus, local $\mathcal{D}(H)$ gauge symmetry cannot be thought as a fixed system of pairwise commuting $\mathcal{D}(H)$-actions. Although there are maximal systems of such actions, they are not unique and may not involve all F and G operators, especially if $|\Sigma^0| \neq |\Sigma^2|$.

There is also another pair of left and right actions,

$$(\psi \otimes k) \underset{a}{\blacktriangleright} M := \mathbf{D}_a\big((\psi \otimes k)_{(2)}\big) M \, \mathbf{D}_a\big(S_{\mathcal{D}}\big((\psi \otimes k)_{(1)}\big)\big)$$
$$= F_a(\psi_{(1)}) G_a(k_{(2)}) \, M \, G_a\big(S(k_{(1)})\big) F_a\big(S(\psi_{(2)})\big),$$
$$M \underset{a}{\blacktriangleleft} (\psi \otimes k) := \mathbf{D}_a\big(S_{\mathcal{D}}\big((\psi \otimes k)_{(2)}\big)\big) M \, \mathbf{D}_a\big((\psi \otimes k)_{(1)}\big)$$
$$= G_a\big(S(k_{(2)})\big) F_a\big(S(\psi_{(1)})\big) \, M \, F_a(\psi_{(2)}) G_a(k_{(1)})$$

with respect to which, however, not $\mathcal{M}$ but $\mathcal{M}^{\mathrm{op}}$ is a module algebra.

It is worth investigating the subspaces $P_a(H)$ and $Q_a(H^*)$ because they are the smallest $D(H)$-invariant subspaces, in fact module subalgebras, of $\mathcal{M}$ or $\mathcal{M}^{\mathrm{op}}$. The exchange relations

$$G_a(k) Q_a(\varphi) = Q_a\big(\varphi \leftharpoonup S(k_{(2)})\big) G_a(k_{(1)}), \tag{3.23}$$

$$F_a(\psi) Q_a(\varphi) = Q_a\big(\psi_{(1)} \varphi S(\psi_{(2)})\big) F_a(\psi_{(3)}), \tag{3.24}$$

$$Q_a(\varphi) G_{T_2 a}(k) = G_{T_2 a}(k_{(2)}) Q_a\big(S(k_{(1)}) \rightharpoonup \varphi\big), \tag{3.25}$$

$$Q_a(\varphi) F_{T_2 a}(\psi) = F_{T_2 a}(\psi_{(1)}) Q_a\big(S(\psi_{(2)}) \varphi \psi_{(3)}\big) \tag{3.26}$$

follow easily from the defining relations of $\mathcal{M}$, as well as

$$F_{T_0^{-1} a}(\psi) P_a(h) = P_a\big(h \leftharpoonup S(\psi_{(2)})\big) F_{T_0^{-1} a}(\psi_{(1)}), \tag{3.27}$$

$$G_{T_0^{-1} a}(k) P_a(h) = P_a\big(k_{(1)} h S(k_{(2)})\big) G_{T_0^{-1} a}(k_{(3)}), \tag{3.28}$$

$$P_a(h) F_a(\psi) = F_a(\psi_{(2)}) P_a\big(S(\psi_{(1)}) \rightharpoonup h\big), \tag{3.29}$$

$$P_a(h) G_a(k) = G_a(k_{(1)}) P_a\big(S(k_{(2)}) h k_{(3)}\big). \tag{3.30}$$

From these relations, one infers that $P_a(H)$ is an invariant subspace under $\triangleright_{T_0^{-1}a}$ and $\triangleleft_a$ while $Q_a(H^*)$ is an invariant subspace under $\blacktriangleright_a$ and $\blacktriangleleft_{T_2a}$. So we have

left $\mathcal{D}(H)$-module subalgebras $\qquad P_a(H) \subset \langle \mathcal{M}, \underset{T_0^{-1}a}{\triangleright} \rangle, \qquad Q_a(H^*)^{\mathrm{op}} \subset \langle \mathcal{M}^{\mathrm{op}}, \underset{a}{\blacktriangleright} \rangle,$

right $\mathcal{D}(H)$-module subalgebras $\qquad P_a(H) \subset \langle \mathcal{M}, \underset{a}{\triangleleft} \rangle, \qquad Q_a(H^*)^{\mathrm{op}} \subset \langle \mathcal{M}^{\mathrm{op}}, \underset{T_2a}{\blacktriangleleft} \rangle.$

Despite of having a left and a right action neither $P_a(H)$ nor $Q_a(H^*)$ is a bimodule over $\mathcal{D}(H)$, the former because $\mathcal{O}_0\big(T_0^{-1}a\big) = \mathcal{O}_0(a)$ and the latter because $\mathcal{O}_2(T_2a) = \mathcal{O}_2(a)$. Denoting these modules by drawing an arrow from the left action to the right action,

$$\mathbf{D}_{T_0^{-1}a} \xrightarrow{P_a(H)} \mathbf{D}_a, \qquad \mathbf{D}_a \xrightarrow{Q_a(H^*)} \mathbf{D}_{T_2a}, \tag{3.31}$$

we see that from the point of view of gauge transformations the P_a and the Q_a seem to belong to different edges and the vertices associated to individual $\mathcal{D}(H)$ actions are not the vertices of Σ, not its faces either, but are in bijection with the arrows themselves. Such a complex is the dual of the double of Σ and will be discussed, among others, in the next two sections.

The Hamiltonian of the model, however, can be introduced without moving to the dual of the double. If we assume that H is semisimple then H has a Haar integral i. The Haar integral being a cocommutative element its iterated coproducts $i_{(1)} \otimes \cdots \otimes i_{(n)}$ are invariant under cyclic permutations. Therefore, $G_a(i)$ does not depend on the starting point of the dual loop, i.e., $G_{T_0a}(i) = G_a(i)$. So we may rename $G_a(i)$ as A_v, where $v = \mathcal{O}_0(a)$. Similarly, H^* also has a Haar integral ι and $F_{T_2a}(\iota) = F_a(\iota)$. Renaming $F_a(\iota)$ as B_f, where $f = \mathcal{O}_2(a)$, we have constructed a system of pairwise commuting projections A_v, for $v \in \Sigma^0$, and B_f, for $f \in \Sigma^2$. The Hamiltonian is the sum

$$\mathbf{H}_{\Sigma,H} = \sum_{v \in \Sigma^0} (\mathbf{1} - A_v) + \sum_{f \in \Sigma^2} (\mathbf{1} - B_f). \tag{3.32}$$

There is a remarkable duality of these models. Replacing H with H^* and Σ with Σ^*, there is an algebra isomorphism $\mathcal{M}(\Sigma, H) \xrightarrow{\sim} \mathcal{M}(H^*, \Sigma^*)$ sending the Hamiltonian (3.32) to $\mathbf{H}_{\Sigma^*,H^*}$. This (non-canonical) isomorphism can be lifted to the ribbon operators, see the duality formula (7.19). The general structure of this "duality" is still to be investigated. In this paper, we will use it only as a technical tool to prove one half of Theorem 7.6.

4   Duals, mirror images and doubles in the arrow presentation

Let $\Sigma :: \langle \mathrm{Arr}, T_0, T_2 \rangle$ be an OCPM. The dual of Σ is easy to formulate in the Σ-language, see (2.9). In the arrow presentation, however, we work with directed edges and there are several natural and less natural ways to assign to an arrow of Σ an arrow of Σ^*. Such problems forces us to speak about, at least, the isomorphisms of arrow presentations. (More general morphisms of OCPMs or of their presentations would require to extend the theory to 3-complexes which we cannot afford here.)

Definition 4.1. An isomorphism $\langle A, S_0, S_2 \rangle \to \langle B, T_0, T_2 \rangle$ of arrow presentations is a bijection $f \colon A \to B$ such that $f \circ S_i = T_i \circ f$ for $i = 0, 2$.

In order to see how this definition translates to the polyhedral maps, we need

Definition 4.2. An isomorphism ϕ of CPMs $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle \xrightarrow{\phi} \langle \Sigma', \dim', \mathrm{Bd}', \mathrm{Cb}' \rangle$ is a bijection $\phi \colon \Sigma \to \Sigma'$ which preserves the dimensions of cells, $\dim' \circ \phi = \dim$, and such that

$\mathrm{Bd}(\phi(x)) = \{\phi(y) \mid y \in \mathrm{Bd}(x)\}$. If $\phi \colon \Sigma \to \Sigma'$ is an isomorphism of CPMs between oriented closed polyhedral maps, then ϕ is called orientation preserving if for all $v \in \Sigma^0$ and $f \in \Sigma^2$ the restrictions of ϕ to the neighbourhoods $\mathrm{Nb}(v)$ and $\mathrm{Nb}(f)$ preserve the cyclic orders given on them by the orientations. Such an orientation preserving isomorphism is called briefly an isomorphism of OCPMs.

Lemma 4.3. *Let $\Sigma :: \langle A, T_0, T_2 \rangle$, $\Sigma' :: \langle A', T_0', T_2' \rangle$ and let $f \colon \langle A, T_0, T_2 \rangle \to \langle A', T_0', T_2' \rangle$ be an isomorphism of arrow presentations. Then $\phi(\mathcal{O}_i(a)) := \mathcal{O}_i'(f(a))$ is an isomorphism of OCPMs. This proves, at least for isomorphisms, functoriality of the construction of OCPMs from arrow presentations described by Theorem* 2.7.

Proof. The isomorphism ϕ is the restriction to the orbits of the direct image $f_* \colon 2^A \to 2^{A'}$ which is the natural lift of f to the power sets. So

$$\phi(\mathcal{O}_i(a)) = \{f \circ T_i^n(a) \mid n \in \mathbb{N}\} = \{T_i'^n(f(a)) \mid n \in \mathbb{N}\} = \mathcal{O}_i'(f(a))$$

holds a priori for $i = 0, 2$ but then also for $i = 1$, i.e., ϕ maps T_i-orbits to T_i'-orbits for all i. Therefore, it preserves the dimension of cells and satisfies $b \in \mathcal{O}_i(a)$ if and only if $f(b) \in \phi(\mathcal{O}_i(f(a)))$. If follows that $\mathcal{O}_{i-1}(b) \cap \mathcal{O}_i(a) \neq \varnothing$ if and only if $\mathcal{O}_{i-1}'(f(b)) \cap \mathcal{O}_i'(f(a)) \neq \varnothing$, so it preserves the boundary and coboundary operations by (2.8), i.e., $\mathrm{Bd}(\phi(s)) = \{\phi(t) \mid t \in \mathrm{Bd}(s)\}$ for all $s \in \Sigma$. Finally, since f commutes with the T_i actions, it is obvious that ϕ preserves the orientation. ∎

After this preparation, we can discuss duals of arrow presentations. They are characterized by producing dual polyhedral maps in the sense of (2.9). Let $P = \langle \mathrm{Arr}, T_0, T_2 \rangle$ be a presentation of an OCPM Σ. A dual $\widetilde{P}$ of P can be obtained by following the schema ⊰ which means that the blue arrow a of Σ is interpreted as a red arrow of Σ^* directed to point from the right face of a to the left face of a. The T_2 of the dual, denoted $\tilde{T}_2$, is rotation around the left face of the red arrow, i.e., around the source vertex of the blue one which is T_0. The $\tilde{T}_0$, however, is obtained by rotating around the source vertex of the red arrow, i.e., around the right face of the blue one which is not T_2 applied to the blue arrow. In its stead $\tilde{T}_0 = T_1 T_2 T_1 = T_0 T_2 T_0^{-1}$. Thus,

$$\widetilde{P} := \langle \mathrm{Arr}, T_0 T_2 T_0^{-1}, T_0 \rangle \tag{4.1}$$

is a presentation producing the dual polyhedral map Σ^*. If we follow the other natural schema ⊰, then we get another presentation

$$\overset{\smile}{P} := \langle \mathrm{Arr}, T_2, T_2^{-1} T_0 T_2 \rangle, \tag{4.2}$$

which should produce the same (or isomorphic) dual map Σ^*. Encouraged by these examples, we may experiment with a simpler presentation of the same sort. Let

$$P^\times := \langle \mathrm{Arr}, T_2, T_0 \rangle.$$

Lemma 4.4. *The bijections $T_i \colon \mathrm{Arr} \to \mathrm{Arr}$ provide isomorphisms of arrow presentations*

$$P^\times \xrightarrow{T_0} \widetilde{P} \xrightarrow{T_1} \overset{\smile}{P} \xrightarrow{T_2} P^\times,$$

$$\overset{\smile}{\widetilde{P}} \xrightarrow{\mathrm{id}} P \xrightarrow{\mathrm{id}} \overset{\sim}{\overset{\smile}{P}},$$

$$\overset{\smile}{\widetilde{P}} \xrightarrow{T_1} P \xrightarrow{T_1} \overset{\sim}{\overset{\smile}{P}}.$$

Proof. Let us prove the first

$$\widetilde{\mathrm{T}}_0 T_0 = T_0 T_2 T_0^{-1} T_0 = T_0 T_2 = T_0 \mathrm{T}_0^\times,$$

$$\widetilde{\mathrm{T}}_2 T_0 = T_0 T_0 = T_0 \mathrm{T}_2^\times.$$

The remaining ones are similarly simple.   ∎

Although all the three duals $\widetilde{\mathrm{P}}$, $\widehat{\mathrm{P}}$ and $\mathrm{P}^\times$ are isomorphic, the first two differs from the third in that

$$\widetilde{\mathrm{T}}_0 \widetilde{\mathrm{T}}_2 = \widehat{\mathrm{T}}_0 \widehat{\mathrm{T}}_2 = T_0 T_2 \qquad \text{while} \quad \mathrm{T}_0^\times \mathrm{T}_2^\times = T_2 T_0.$$

In other words, the arrow-opposite arrow pairs, i.e., the T_1-orbits, of $\widetilde{\mathrm{P}}$ and $\widehat{\mathrm{P}}$ are the same as those of P while in $\mathrm{P}^\times$ they are connected by the transformation $T_2 T_0$ which is non-local in P, it sends an arrow by distance 2 away. Of course, $T_2 T_0$ is also an involution, $T_2 T_0 T_2 T_0 = T_2 T_1 T_0 = T_0^{-1} T_0 = \mathrm{id}$, algebraically as good as T_1, so why is T_1 local and the other is not? We have made a choice, not in Definition 2.1 but in the construction of Σ when we called edges the orbits of $T_1 = T_0 T_2$ and not of $T_2 T_0$. The form of $\mathrm{P}^\times$ shows that if we had chosen edges to be the $T_2 T_0$-orbits the interpretation of T_0, T_2 would have changed to T_0 rotating around faces and T_2 rotating around vertices. (Speaking about non-locality of a presentation Q is meaningful only with respect to another isomorphic presentation P: Q is non-local if the isomorphism $P \to Q$ is non-local, i.e., tears apart the arrows of an edge. Of course, both T_0 and T_2 are non-local and T_1 is local in this sense.)

When we draw a portion of a polyhedral map we look at it from a specific side of 3-space determined by the orientation. If we look at it from the other side what we see is the mirror image. In more precise terms, if P is a presentation of an OCPM Σ, then the mirror image should be the presentation of the same underlying CPM but equipped with opposite orientation. It is easy to see that the mirror image of $P = \langle \mathrm{Arr}, T_0, T_2 \rangle$ can be presented by

$$P^{\mathbf{m}} := \langle \mathrm{Arr}, T_0^{-1}, T_1 T_2^{-1} T_1 \rangle = \langle \mathrm{Arr}, T_0^{-1}, T_0 T_2^{-1} T_0^{-1} \rangle. \tag{4.3}$$

Another, simpler but non-local, presentation is $P^{\mathbf{M}} := \langle \mathrm{Arr}, T_0^{-1}, T_2^{-1} \rangle$ with isomorphism $T_0 \colon P^{\mathbf{M}} \longrightarrow P^{\mathbf{m}}$.

In order to get an arrow presentation for the double $\mathcal{D}(\Sigma)$, we first have to enumerate the arrows of the double in some way. From Definition 2.9, one infers that $\mathcal{D}(\Sigma)^1$ is 4 times bigger than Σ^1 and the same should hold for the arrows. Let $e = \mathcal{O}_1(a) \in \Sigma^1$ be considered as a vertex of the double. There are 4 edges joining at e, 2 ve-type edges and 2 ef-type edges. We denote the arrow $\mathcal{O}_0(a) \to e$ by a_0^+ and the arrow $\mathcal{O}_2(a) \to e$ by a_2^-. Their opposites are called a_0^- and a_2^-, respectively. The arrows of the remaining two edges joining to e can then be expressed as $(T_1 a)_0^\pm$ and $(T_1 a)_2^\pm$. The identities

$$\mathcal{O}_0(a) \cap \mathcal{O}_1(T_0 a) = \mathcal{O}_0(T_0 a) \cap \mathcal{O}_1(T_0 a) = \{T_0 a\},$$

$$\mathcal{O}_2(a) \cap \mathcal{O}_1(T_2^{-1} a) = \mathcal{O}_2(T_2^{-1} a) \cap \mathcal{O}_1(T_2^{-1} a) = \{T_2^{-1} a\}$$

show that the arrows and vertices of every face neighbourhood of the double can be labelled by
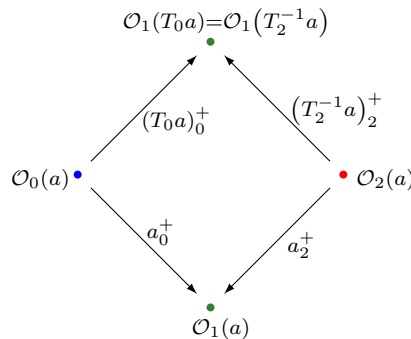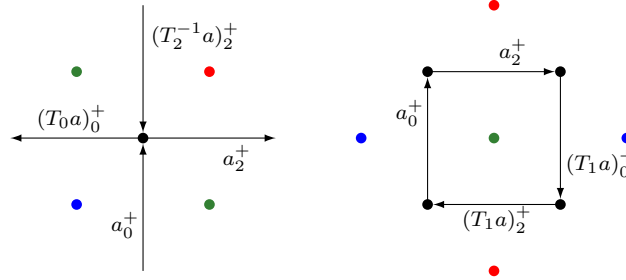
Figure 2. The vertex neighbourhood of the site $\langle \mathcal{O}_0(a), \mathcal{O}_2(a) \rangle$ and the face neighbourhood of the type 1 face $\mathcal{O}_1(a)$.

Thus, the arrows of $\mathcal{D}(\Sigma)$ is the set $\mathbb{A} := \{a_i^\sigma \mid a \in \mathrm{Arr}, \, i \in \{0,2\}, \, \sigma \in \{+,-\}\}$ and the arrow presentation

$$\mathcal{D}(\Sigma) :: \mathcal{D}(P) = \langle \mathbb{A}, \mathbb{T}_0, \mathbb{T}_2 \rangle$$

is given by the following permutations:

$$
\mathbb{T}_0 a_i^\sigma := \begin{cases} (T_0 a)_0^+ & \text{if } i = 0, \, \sigma = +, \\ (T_1 a)_2^- & \text{if } i = 0, \, \sigma = -, \\ (T_2 a)_2^+ & \text{if } i = 2, \, \sigma = +, \\ a_0^- & \text{if } i = 2, \, \sigma = -, \end{cases}
\qquad
\mathbb{T}_2 a_i^\sigma := \begin{cases} a_2^- & \text{if } i = 0, \, \sigma = +, \\ (T_0^{-1} a)_0^+ & \text{if } i = 0, \, \sigma = -, \\ (T_1 a)_0^- & \text{if } i = 2, \, \sigma = +, \\ (T_2^{-1} a)_2^+ & \text{if } i = 2, \, \sigma = -. \end{cases}
\tag{4.4}
$$

Defining also $\mathbb{T}_1 = \mathbb{T}_0 \mathbb{T}_2$, we obtain

$$\mathbb{T}_1 a_i^\sigma = a_i^{-\sigma}, \qquad \mathbb{T}_2^4 = \mathrm{id}_\mathbb{A}, \qquad \mathbb{T}_0^4 a_i^- = a_i^-$$

for all a, i and σ. The second expresses the fact that all faces of the double are quadrangles and the third holds (only for $\sigma = -$) because the e-type vertices of $\mathcal{D}(\Sigma)$ all have degree 4.

Roughly speaking, the T-operators of the double $\mathcal{D}(\Sigma)$ are matrix amplifications of the T-operations of Σ,

$$
\mathbb{T}_0 = \begin{pmatrix} T_0 & & & \\ & & & 1 \\ & & T_2 & \\ & T_1 & & \end{pmatrix},
\qquad
\mathbb{T}_2 = \begin{pmatrix} & T_0^{-1} & & \\ & & T_1 & \\ & & & T_2^{-1} \\ 1 & & & \end{pmatrix},
$$

where the order of rows and columns is that of formulas (4.4) and 1 stands for $\mathrm{id}_{\mathrm{Arr}}$. But these are not matrices in the usual sense: The addition involved in matrix multiplications is not defined at all. (We are working in the category of sets after all.) This is why we have left the nonvalid entries empty rather than put a zero there.

Having been constructed arrow presentations for Σ^* and $\mathcal{D}(\Sigma)$ of every OCPM Σ has the secondary role of extending the Definition 2.9 by giving them orientations.

Now it is easy to obtain an arrow presentation for the dual of the double of Σ. We take the above presentation of $\mathcal{D}(\Sigma)$ and then apply the construction $\tilde{\mathrm{P}}$. This leads to the following presentation for $\mathcal{D}(\Sigma)^*$. The set of arrows is $\{a_i^\sigma \mid i = 0, e, \, \sigma = \pm\}$, the same as that of the double. (We spare the work of writing $\tilde{a}_i^\sigma$ to distinguish them from the arrows of $\mathcal{D}(\Sigma)$, eventually it is the same set.) These arrows are shown on Figure 2. Thus, we have the presentation

$$\mathcal{D}(\Sigma)^* :: \mathcal{D}(P)^\sim = \langle \mathbb{A}, \tilde{\mathbb{T}}_0, \tilde{\mathbb{T}}_2 \rangle \tag{4.5}$$

with

$$\tilde{\mathbb{T}}_0 = \begin{pmatrix} & & T_1 \\ T_0^{-1} & & \\ & 1 & \\ & & T_2^{-1} \end{pmatrix}, \qquad \tilde{\mathbb{T}}_2 = \begin{pmatrix} T_0 & & \\ & & 1 \\ & T_2 & \\ T_1 & & \end{pmatrix}.$$

The set $\mathcal{D}(\Sigma)^{*2}$ of faces of the dual of the double is the same set as the set Σ of all cells of the original OCPM. Changing between these two interpretations of the same set will be a recurring tool in this paper. Since the role of these faces depend on their dimensions in Σ and we want to avoid saying "vertex face" or "face face" we shall use the term "type d face" for an element $c \in \Sigma^d \subset \mathcal{D}(\Sigma)^{*2}$.

Proposition 4.5. *There exist natural isomorphisms of arrow presentations*

$$\begin{aligned} \delta_P\colon\ & \mathcal{D}(P)^\sim \to \mathcal{D}(P^\sim)^\sim, & a_0^\sigma &\mapsto a_2^\sigma, & a_2^\sigma &\mapsto (T_1 a)_0^\sigma, \\ \mu_P\colon\ & \mathcal{D}(P^{\mathbf{m}})^\sim \to (\mathcal{D}(P)^\sim)^{\mathbf{m}}, & a_0^\sigma &\mapsto a_0^{-\sigma}, & a_2^\sigma &\mapsto (T_1 a)_2^{-\sigma} \end{aligned}$$

or in matrix form

$$\delta_P = \begin{pmatrix} & T_1 & \\ & & T_1 \\ 1 & & \\ & 1 & \end{pmatrix}, \qquad \mu_P = \begin{pmatrix} & 1 & \\ 1 & & \\ & & T_1 \\ & T_1 & \end{pmatrix}.$$

Proof. Collect the definitions (4.1), (4.5) and (4.3) and perform the matrix multiplications. ∎

5   Curves, dual curves and ribbons

Curves on surfaces do not require orientation of the surface. Nevertheless, we assume that Σ is an OCPM equipped with arrow presentation $P = \langle \mathrm{Arr}(\Sigma), T_0, T_2 \rangle$. Both curves and opposite curves (see opcurves below) on Σ formulate the idea of what is called a walk on a directed graph. By directed graph, we mean a quiver

$$A \underset{\text{target}}{\overset{\text{source}}{\rightrightarrows}} V$$

consisting of two sets, two functions and no axioms. Although the CPM $\langle \Sigma, \dim, \mathrm{Bd}, \mathrm{Cb} \rangle$ has an underlying graph, it is undirected. It consists of the sets Σ^1 and Σ^0 and of the function Bd which associates to every edge $e \in \Sigma^1$ the 2-element subset $\mathrm{Bd}(e) \subseteq \Sigma^0$. Since this undirected graph has no loops, there is a natural way to make it a quiver by doubling the set of edges. Making use of the arrow presentation P, we can take the set $\mathrm{Arr}(\Sigma)$ of arrows and define the quiver

$$\mathrm{Qv}(P) := \left(\mathrm{Arr}(\Sigma) \underset{\partial_1}{\overset{\partial_0}{\rightrightarrows}} \Sigma^0 \right) \qquad \text{where} \quad \partial_0 a := \mathcal{O}_0(a), \quad \partial_1 a := \mathcal{O}_0(T_1 a).$$

This quiver is involutive in the sense of having an involution T_1 such that $\partial_0 \circ T_1 = \partial_1$; this way reminding us that it stems from an undirected graph.

The difference between curves and opcurves is merely conventional; the literature on Kitaev models seems to use both curves [3, 15] and opcurves [1]. In this paper, most of the calculations will be done with opcurves. The reason is that opcurves fit better with the conventions we have chosen in Definition 2.1.

Definition 5.1.

(a) An "opcurve on Σ", more precisely a P-opcurve, is a sequence $\gamma = (a_1, \ldots, a_n)$ of arrows such that $\partial_1 a_i = \partial_0 a_{i+1}$ for $0 < i < n$. If $\partial_0 a_1 = v_0$ and $\partial_1 a_n = v_n$, then we say that γ is an opcurve from v_0 to v_n and write $\gamma \colon v_0 \to v_n$. The opcurves on Σ form a category with composition of opcurves $u \xrightarrow{\alpha} v \xrightarrow{\beta} w$ being the concatenation $\alpha\beta \colon u \to w$.

(b) A "curve on Σ", i.e., a P-curve, is a sequence $\gamma = (a_1, \ldots, a_n)$ of arrows such that the reverse sequence $\gamma^{\mathrm{rev}} = (a_n, \ldots, a_1)$ is an opcurve, that is to say, such that $\partial_0 a_i = \partial_1 a_{i+1}$ for $0 < i < n$. This curve has source $\partial_0 a_n$ and target $\partial_1 a_1$. The curves on Σ form a category with composition of curves $u \xrightarrow{\alpha} v \xrightarrow{\beta} w$ being $\beta \circ \alpha = \alpha\beta \colon u \to w$. So this category is just the opposite of the category of opcurves.

(c) If an opcurve $\gamma = (a_1, \ldots, a_n)$ is open, i.e., $\partial_0 a_1 \neq \partial_1 a_n$, then it is called simple if the sequence $(\partial_0 a_1, \partial_1 a_1, \partial_1 a_2, \ldots \partial_1 a_n)$ of visited vertices contains no repetitions. If γ is closed, i.e., $\partial_0 a_1 = \partial_1 a_n$, then it is called simple if $(\partial_1 a_1, \partial_1 a_2, \ldots \partial_1 a_n)$ contains no repeated vertices. A curve γ is called simple if the opcurve γ^{rev} is simple.

It is clear that passing to the reverse sequence provides a bijection between curves and opcurves. Another important operation is the "inverse" curve: If $\gamma = (a_1, \ldots, a_n)$ is a curve (opcurve), then $\gamma^{-1} := (T_1 a_n, \ldots, T_1 a_1)$ is also a curve (opcurve). But it is not the inverse of γ until we pass to the paths of the (op)curve. Paths are equivalence classes of curves with respect to insertion and deletion of length 2 subcurves of the form $(a, T_1 a)$. Every equivalence class contains a unique element which does not contain any $(a, T_1 a)$ subcurves, i.e., which is reduced. We shall not introduce any special notation for paths, just say "the path γ" for a curve γ if we think of its equivalence class. This convention is used also for opcurves, so we say "the path γ" even if γ is an opcurve. The paths of curves, as well as the paths of opcurves, form a groupoid $\mathsf{Path}(\Sigma)$ and $\mathsf{Path}(\Sigma)^{\mathrm{op}}$, respectively. Of course, this path groupoid is not the fundamental groupoid of the surface $[\Sigma]$, it is that of the 1-skeleton of Σ, so this groupoid sees every face of Σ as a hole on $[\Sigma]$.

Although the definition of (op)curves and therefore also of the path groupoid depend on the arrow presentation, one can easily verify that any isomorphism of arrow presentations induces an isomorphism between the corresponding categories of (op)curves as well as an isomorphism of their path groupoids.

Using the term P-curve instead of "curve on Σ" is nevertheless advisable, especially if we want to define dual curves. The dual (op)curves on Σ can be thought as (op)curves on the dual Σ^*. But if we want to see them as sequences of arrows in Σ then we need a mapping $\mathrm{Arr}(\Sigma^*) \to \mathrm{Arr}(\Sigma)$. Such mappings are not unique as it is witnessed by the two duals $\widetilde{P}$ and $\breve{P}$ of a presentation. So, if the context requires the precision, we shall speak about $\widetilde{P}$-(op)curves and $\breve{P}$-(op)curves instead of "dual (op)curves on Σ".

Definition 5.2. A $\widetilde{P}$-opcurve is a sequence $\beta = (a_1, \ldots, a_n)$ of arrows $a_i \in \mathrm{Arr}(\Sigma)$ such that $d_0 a_i = d_1 a_{i+1}$ for $0 < i < n$, where we introduced the notation

$$d_0 a := \mathcal{O}_2(a), \qquad d_1 a := \mathcal{O}_2(T_1 a).$$

The $\breve{P}$-opcurves are defined as the reverses of $\widetilde{P}$-opcurves.

These definitions are based on the quivers

$$\mathrm{Qv}(\widetilde{P}) = \left(\mathrm{Arr}(\Sigma) \underset{d_0}{\overset{d_1}{\rightrightarrows}} \Sigma^0 \right), \qquad \mathrm{Qv}(\breve{P}) = \left(\mathrm{Arr}(\Sigma) \underset{d_1}{\overset{d_0}{\rightrightarrows}} \Sigma^0 \right).$$

In fact, the $\widetilde{P}$-opcurves coincide with the $\breve{P}$-curves and vice versa.

Remark 5.3. Passing from curves to opcurves, or back, can be automatized by introducing $P^{\mathrm{op}} := \langle \mathrm{Arr}(\Sigma), T_1 T_0 T_1, T_1 T_2 T_1 \rangle$ which is a presentation isomorphic to P via the transformation T_1. Since T_1 is also the involution of the quiver, so

$$\mathrm{Qv}(P^{\mathrm{op}}) = \left(\mathrm{Arr}(\Sigma) \underset{\partial_0}{\overset{\partial_1}{\rightrightarrows}} \Sigma^0 \right),$$

and therefore the P^{op}-curves are the P-opcurves. Also, we have $\left(\widetilde{P} \right)^{\mathrm{op}} = \widecheck{P}$. Although this might suggest a redundancy of using both curves and opcurves but the concept of P-opcurve comprises not only the use of P^{op}-curves but also the convention of using concatenation for their composition.

Now we turn to the description of curves on the dual of the double $\mathcal{D}(\Sigma)^*$ of an OCPM Σ. The dual of the double is particularly interesting for two reasons: Geometrically because it provides the Schreier coset graph for the arrow presentation, and physically because it is the surface complex on which holonomy of the Kitaev model can be defined.

Fixing an arrow presentation $P = \langle A, T_0, T_2 \rangle$ for Σ, the vertices of $\mathcal{D}(\Sigma)^*$ can be identified with the set of sites, i.e., with pairs $\langle v, f \rangle \in \Sigma^0 \times \Sigma^2$ such that $v \cap f \neq \varnothing$, cf. (2.4) and Lemma 2.4. The set of arrows of $\mathcal{D}(\Sigma)^*$ can be identified with the set $\mathbb{A} = \{ a_i^\sigma \mid a \in A, i = 0, 2, \sigma = \pm \}$, see Figure 2. The meaning of the arrows can be made explicit by giving the source and target maps $\nabla_0, \nabla_1 \colon \mathbb{A} \to \mathcal{D}(\Sigma)^{*0}$,

$$\begin{aligned}
\nabla_0 a_0^+ &= \langle \partial_0 a, d_1 a \rangle \\
\nabla_1 a_0^+ &= \langle \partial_0 a, d_0 a \rangle \\
\nabla_0 a_2^+ &= \langle \partial_0 a, d_0 a \rangle \\
\nabla_1 a_2^+ &= \langle \partial_1 a, d_0 a \rangle
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\nabla_0 a_i^- &= \nabla_1 a_i^+ \\
\nabla_1 a_i^- &= \nabla_0 a_i^+.
\end{aligned}
\tag{5.1}$$

The quiver $\mathbb{A} \underset{\nabla_1}{\overset{\nabla_0}{\rightrightarrows}} \mathcal{D}(\Sigma)^{*0}$ can be seen to be the quiver associated to the arrow presentation (4.5) so we call it $\mathrm{Qv}(\mathcal{D}(P)^\sim)$. Curves and opcurves on $\mathcal{D}(\Sigma)^*$ will always be meant with respect to this quiver.

Lemma 5.4. *The following map embeds the involutive quiver of $\mathcal{D}(P)^\sim$ to the cartesian product of the involutive quivers of P and $\widetilde{P}$:*

$$\left(\mathbb{A} \underset{\nabla_1}{\overset{\nabla_0}{\rightrightarrows}} \mathcal{D}(\Sigma)^{*0} \right) \overset{E}{\longrightarrow} \left(A \underset{\partial_1}{\overset{\partial_0}{\rightrightarrows}} \Sigma^0 \right) \times \left(A \underset{d_0}{\overset{d_1}{\rightrightarrows}} \Sigma^2 \right),$$

$$\begin{aligned}
\langle v, f \rangle &\mapsto \langle v, f \rangle, \\
a_0^+ &\mapsto \langle \partial_0 a, a \rangle, \\
a_0^- &\mapsto \langle \partial_0 a, T_1 a \rangle, \\
a_2^+ &\mapsto \langle a, d_0 a \rangle, \\
a_2^- &\mapsto \langle T_1 a, d_0 a \rangle.
\end{aligned}$$

Proof. Recall that the cartesian product of quivers $\mathbb{Q}$ and $\mathbb{Q}'$ has set of vertices $(\mathbb{Q} \times \mathbb{Q}')^0 = \mathbb{Q}^0 \times \mathbb{Q}'^0$ and set of arrows $(\mathbb{Q} \times \mathbb{Q}')^1 = \left(\mathbb{Q}^0 \times \mathbb{Q}'^1 \right) \sqcup \left(\mathbb{Q}^1 \times \mathbb{Q}'^0 \right)$ and it is endowed with the obvious source and target maps. In case of the quivers of Σ and Σ^*, these obvious maps are $\partial_0 \times d_1$ and $\partial_1 \times d_0$ with the warning that here the ∂_i and d_i are used in a generalized sense: The ∂_i is extended to the trivial edges of Σ, i.e., to the vertices, by $\partial_0 v = \partial_1 v = v$ and the d_i is extended to the trivial edges of Σ^*, i.e., to the faces, by $d_0 f = d_1 f = f$. Then the fact that E is a homomorphism of involutive quivers,

$$(\partial_0 \times d_1) \circ E = E \circ \nabla_0, \qquad (\partial_1 \times d_0) \circ E = E \circ \nabla_1, \qquad (T_1 \times \tilde{T}_1) \circ E = E \circ \tilde{\mathbb{T}}_1$$

can be easily verified. That E is injective is also clear. ∎
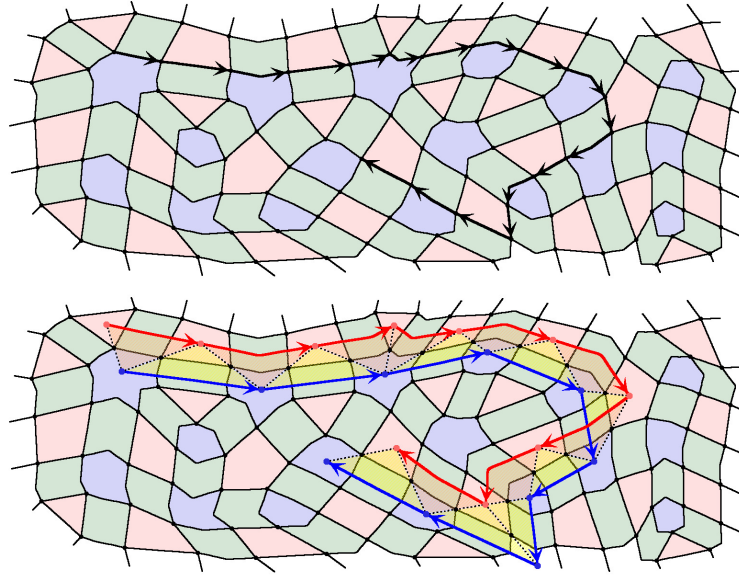
Figure 3. Every curve ρ (black) on $\mathcal{D}(\Sigma)^*$ is accompanied by a curve γ (blue) on Σ and a curve β (red) on Σ^* according to Lemma 5.4. They border a strip made of triangles but the strip may fail to be embedded into the surface of Σ: The γ and β intersect locally, i.e., the strip folds, precisely if the ρ makes a turn around a type 1 (green) face of $\mathcal{D}(\Sigma)^*$.

In the image of E the arrows can be visualized as "isosceles triangles". The triangle $E\big(a_2^\pm\big)$ has base edge a with opposite vertex being the face $f = d_0 a$ and the legs of the triangle being the sites $\langle v, f \rangle$ where v is either $\partial_0 a$ or $\partial_1 a$. The triangles $E\big(a_0^\pm\big)$ are called dual triangles because in them the roles of vertices and faces are interchanged. If $\rho = (d_j)$ is a curve on $\mathcal{D}(\Sigma)^*$, then two consecutive triangles $E(d_j)$ and $E(d_{j+1})$ are pasted together at a leg so the whole picture of the curve looks like a ribbon with the base edges of triangles forming one borderline and that of the dual triangles (drawn as edges of Σ^*) forming the other borderline. On the two ends, we see two unpaired legs of some triangles and the texture of the ribbon itself is a triangulated surface, see Figure 3. The 2 borderline curves can be obtained by writing $E(d_j)$ as $\langle c_j, b_j \rangle$ and then $\gamma = (c_j)$ is a P-curve and $\beta = (b_j)$ is a $\widetilde{\mathrm{P}}$-curve in the generalized sense: γ may contain vertices and β may contain faces. But c_j is a vertex if and only if b_j is a genuine arrow and c_j is a genuine arrow if and only if b_j is a face, hence the pair $\langle c_j, b_j \rangle$ is always a triangle.

This is the origin of the names such as "triangle operators" and "ribbon curves" in the literature on the Kitaev model, although the name ribbon curve is usually reserved for a special kind of curve on $\mathcal{D}(\Sigma)^*$ which we will define below.

In the next lemma, we take a closer look into the structure of $\mathcal{D}(\Sigma)^*$. We will see that we can forget about the complicated arrow presentation (4.5) and navigate through $\mathcal{D}(\Sigma)^*$ using essentially only the arrow presentation of Σ.

Lemma 5.5. *The map $a \mapsto s(a) := \langle \mathcal{O}_0(a), \mathcal{O}_2(a) \rangle$ is a bijection $\mathrm{Arr}(\Sigma) \cong \mathcal{D}(\Sigma)^{*0}$. Every vertex of $\mathcal{D}(\Sigma)^*$ has degree 4 and the four arrows pointing out from $s(a)$ can be labelled, or colored, by the letters $\big\{T_0, T_0^{-1}, T_2, T_2^{-1}\big\}$ as follows*

$$\begin{aligned}
(T_0 a)_0^+ : & \quad s(a) \to s(T_0 a) & & \textit{labelled by } T_0, \\
a_0^- : & \quad s(a) \to s\big(T_0^{-1} a\big) & & \textit{labelled by } T_0^{-1}, \\
a_2^+ : & \quad s(a) \to s(T_2 a) & & \textit{labelled by } T_2, \\
\big(T_2^{-1} a\big)_2^- : & \quad s(a) \to s\big(T_2^{-1} a\big) & & \textit{labelled by } T_2^{-1} \, .
\end{aligned}$$

The order in which they are listed corresponds to the cyclic order induced by the action of $\tilde{\mathbb{T}}_0$ of (4.5). Moreover, if an arrow d carries the label T_j^ϵ then $\tilde{\mathbb{T}}_1 d$ has label $T_j^{-\epsilon}$.

Proof. $\mathcal{O}_0(a) \cap \mathcal{O}_2(a) = \{a\}$ is non-empty, so $s(a)$ is a site, cf. Definition 2.5 and (2.4). If $\langle v, f \rangle$ is a site, then $v \cap f \neq \varnothing$, so $v = \mathcal{O}_0(a)$ and $f = \mathcal{O}_2(a)$ for the unique element $a \in v \cap f$. So s is indeed a bijection.

The four arrows of source $s(a)$ are dual to the 4 boundary arrows of the rhombus face $s(a)$ of the double $\mathcal{D}(\Sigma)$. (The reason for drawing $\mathcal{D}(\Sigma)$ and not $\mathcal{D}(\Sigma)^*$ is merely typographical.)

$$
\begin{aligned}
v &= \mathcal{O}_0(a), & f &= \mathcal{O}_2(a) \\
e^L &= \mathcal{O}_1(T_0 a), & e^R &= \mathcal{O}_1(a) \\
v^L &= \mathcal{O}_0(T_2^{-1} a), & v^R &= \mathcal{O}_0(T_1 a) \\
f^L &= \mathcal{O}_2(T_0 a), & f^R &= \mathcal{O}_2(T_1 a).
\end{aligned}
\tag{5.2}
$$

The e^L, v, e^R, f are the 4 vertices of the rhombus by (2.5), $\{v, v^L\} = \mathrm{Bd}(e^L)$, $\{v, v^R\} = \mathrm{Bd}(e^R)$, $\{f, f^L\} = \mathrm{Cb}(e^L)$, $\{f, f^R\} = \mathrm{Cb}(e^R)$ and the four neighbour sites of $\langle v, f \rangle$ are $\langle v, f^L \rangle$, $\langle v, f^R \rangle$, $\langle v^R, f \rangle$ and $\langle v^L, f \rangle$. Computing the intersections of the vertex and face within these sites, we recognize them as $s(T_0 a)$, $s(T_0^{-1} a)$, $s(T_2 a)$ and $s(T_2^{-1} a)$, respectively. (For example, $v \cap f^L = \mathcal{O}_0(a) \cap \mathcal{O}_2(T_0 a) = \mathcal{O}_0(T_0 a) \cap \mathcal{O}_2(T_0 a) = \{T_0 a\}$.) This partially explains the labels of the four arrows, partially because some of these 4 sites can be identical. (Indeed, $\mathcal{D}(\Sigma)^*$ may have double arrows but never multiple arrows of larger multiplicity.) $s(T_0 a) = s(T_0^{-1} a)$ occurs when v has degree 2 and $s(T_2 a) = s(T_2^{-1} a)$ occurs when f is a 2-gon. In these degenerate cases

there is a double arrow from $s(a)$ to, let's say, $s(T_2 a)$ (the 2nd picture above) and it seems to be indifferent which is labelled by T_2 because T_2 acts on a as T_2^{-1}. However, the letters T_2 and T_2^{-1} are different and so are the two parallel arrows. The labelling in this case is defined by forcing the lemma: The labels must follow the cyclic order $[T_0, T_0^{-1}, T_2, T_2^{-1}]$ when we apply $\tilde{\mathbb{T}}_0$. There is only one case left. If both types of degeneracy occur for the same arrow a, then $T_0^2 a = a = T_2^2 a$ so by Example 2.10 the connected component of a is nothing but the minimal OCPM, $\Sigma_{\min}$. In this case, the correctness of the labelling can be checked by inspection. ∎

The above lemma has two important consequences. The first one is an interpretation of $\mathcal{D}(\Sigma)^*$ as the coset graph of the arrow presentation of Σ. This is the next corollary. The second one is a useful encoding of the curves on $\mathcal{D}(\Sigma)^*$ in terms of words in the alphabet $\{T_0, T_0^{-1}, T_2, T_2^{-1}\}$, see the definition below.

Corollary 5.6. *Let Σ be a connected OCPM. Every arrow presentation $\langle \mathrm{Arr}, T_0, T_2 \rangle$ of Σ determines a transitive action on $\mathrm{Arr}(\Sigma)$ of the Σ-independent group $\mathcal{T} = \langle T_0, T_2 \mid T_0 T_2 T_0 T_2 \rangle$ and therefore the arrow presentation can be thought as a presentation of the left coset space $\mathcal{T}/\mathcal{H}$ of the group $\mathcal{T}$ with respect to some subgroup. Then the Schreier coset graph of $\mathcal{T}/\mathcal{H}$ can be drawn on the surface $[\Sigma]$ as follows. Each $a \in \mathrm{Arr}(\Sigma)$ is drawn as a point at $s(a) \in \mathcal{D}(\Sigma)^{*0}$. For each $a \in \mathrm{Arr}(\Sigma)$, we draw two arrows: $(T_0 a)_0^+ : s(a) \to s(T_0 a)$ colored blue and $a_2^+ : s(a) \to s(T_2 a)$*
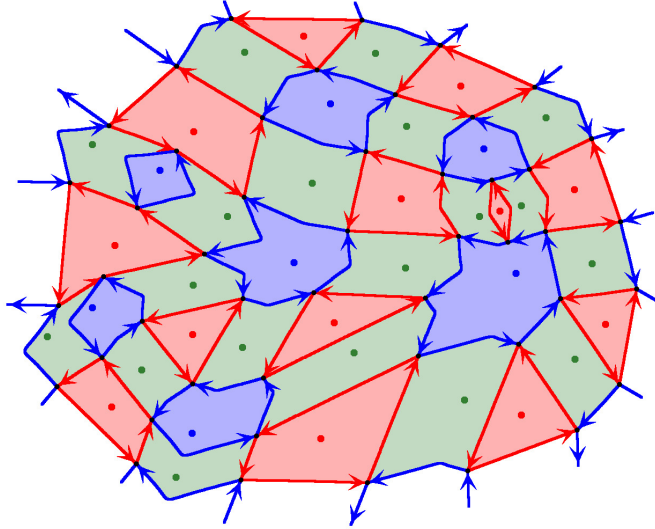
Figure 4. The dual of the double of an OCPM Σ provides the Schreier coset graph of the arrow presentation of Σ. The blue and red arrows show the action of T_0 and T_2, respectively. The green quadrangles represent the universal relation $T_0 T_2 T_0 T_2 = \mathrm{id}$.

colored red. Then every edge of $\mathcal{D}(\Sigma)^*$ is drawn exactly once. The blue arrows represent the flow of the permutation T_0 and the red ones represent that of T_2. In this sense drawing the dual of the double we draw the arrow presentation of Σ (Figure 4).

Definition 5.7. Using the function $s\colon \mathrm{Arr}(\Sigma) \to \mathcal{D}(\Sigma)^{*0}$ defined in the above lemma a curve $\rho = (d_n, \ldots, d_1)$ on $\mathcal{D}(\Sigma)^*$ can be encoded by the pair

$$\langle a_0, w \rangle \in \mathrm{Arr}(\Sigma) \times \left\{ T_0, T_0^{-1}, T_2, T_2^{-1} \right\}^*,$$

where $s(a_0) = \nabla_0 d_1$ is the source vertex of the curve, $w = W_n \cdots W_1$ is the code word made of letters $W_j \in \left\{ T_0, T_0^{-1}, T_2, T_2^{-1} \right\}$ in such a way that for $0 < j \leq n$ the d_j is the unique arrow of source $s(a_{j-1})$ and label W_j, so its target can be written as $s(a_j)$ with $a_j = W_j a_{j-1}$.

The same pair $\langle a_0, w \rangle$ is used to encode the opcurve $\rho^{\mathrm{rev}} = (d_1, \ldots, d_n)$.

Now we are in a position to define ribbons.

Lemma 5.8. *Let $\rho = (d_1, \ldots, d_n)$ be either a curve or an opcurve on $\mathcal{D}(\Sigma)^*$ with encoding $\langle a_0, w \rangle$. The following conditions are equivalent:*

(i) *The triangles $E(d_j)$ and $E(d_{j+1})$ are non-overlapping (their interiors are disjoint) for $0 < j < n$.*

(ii) *ρ is reduced and the 2 borderlines γ and β of the "ribbon" ρ do not intersect locally, i.e., $\mathcal{O}_1(b_j) \neq \mathcal{O}_1(c_{j\pm1})$ whenever both b_j and $c_{j\pm1}$ are non-trivial arrows.*

(iii) *ρ is reduced and if (d_j, d_{j+1}) is a left or a right turn (not straight), then the unique face it turns around is either a type 0 face or a type 2 face, never a type 1 face.*

(iv) *The word w is reduced and does not contain $T_0 T_2$, $T_2 T_0$ and their inverses.*

(v) *The word w belongs to either $\left\{ T_0^{-1}, T_2 \right\}^*$ or $\left\{ T_0, T_2^{-1} \right\}^*$.*

Proof. Let $\langle v, f \rangle = \nabla_0 d_{j+1} = \nabla_1 d_j$. Then the triangles $E(d_j)$ and $E(d_{j+1})$ contain $\langle v, f \rangle$ as a leg. There are 4 such triangles. In the notations of (5.2), they are vfv^L, fvf^L, vfv^R and

fvf^R. The first two and the second two overlap because their base edges intersect at their "middle" points e^L and e^R, respectively. This intersection is, at the same time, an intersection of the two borderlines γ and β of the ribbon. No other pairs of triangles are overlapping except when the two triangles are the same, this happens precisely when $d_{j+1} = \tilde{\mathbb{T}}_1 d_j$.

In case of intersection at e^L the (d_j, d_{j+1}) is either $\left((T_2^{-1}a)_2^+, (T_0a)_0^+\right)$ or its inverse and the corresponding code word is T_0T_2 or its inverse. In case of intersection at e^R, the (d_j, d_{j+1}) is either $\left(a_0^+, a_2^+\right)$ or its inverse and the corresponding code word is T_2T_0 or its inverse. These pairs of arrows make turn around the type 1 face e^L and e^R, respectively. All other pairs of arrows are either straight or make a turn around v or f.

If a reduced word w does not contain the alphabet changing words T_0T_2, T_2T_0 and their inverses, then $W_1 \in \left\{T_0^{-1}, T_2\right\}$ is equivalent to $w \in \left\{T_0^{-1}, T_2\right\}^*$. The same holds for the alphabet $\left\{T_0, T_2^{-1}\right\}$. ∎

Definition 5.9. A ribbon (op)curve is a (op)curve on $\mathcal{D}(\Sigma)^*$ satisfying the conditions of Lemma 5.8. Let $\mathsf{Ribb}(\Sigma)$ and $\mathsf{Ribb}^\circ(\Sigma)$ be the set of ribbon curves and ribbon opcurves, respectively. If the code word of a ribbon (op)curve belongs to $\left\{T_0^{-1}, T_2\right\}^*$, then it is called a left ribbon (op)curve. If it belongs to $\left\{T_0, T_2^{-1}\right\}^*$, we speak about right ribbon (op)curve. Let $\mathsf{Ribb}_L(\Sigma)$, $\mathsf{Ribb}_R(\Sigma)$, $\mathsf{Ribb}_L^\circ(\Sigma)$, $\mathsf{Ribb}_R^\circ(\Sigma)$ be the set of left/right ribbon curves/opcurves, respectively.

Since every ribbon is reduced, $\mathsf{Ribb}(\Sigma)$ embeds into the path groupoid $\mathsf{Path}(\Sigma)$. But it is not a subgroupoid and not even a subcategory as one can infer from the characterization Lemma 5.8 (v). The $\mathsf{Ribb}_L(\Sigma)$ and $\mathsf{Ribb}_R(\Sigma)$, however, are subcategories with the inverse of every left ribbon being a right ribbon. Moreover, $\mathsf{Ribb}(\Sigma) = \mathsf{Ribb}_L(\Sigma) \cup \mathsf{Ribb}_R(\Sigma)$. Analogous remarks can be made for the ribbon opcurves.

The above definition of ribbons seems to agree with that of [1] but it is less restrictive than that of [15]. The latter type of ribbons will appear here as proper ribbons in Definition 6.1. The necessity of distinguishing two kinds of ribbons has been realized already in [22], where they are called locally clockwise and locally counterclockwise. Our terminology of left and right refers to that whether the red dual curve on Figure 3 is on the left or on the right of the blue curve. So ribbons may also be described as curves on Σ together with a consistent framing by a dual curve. It is a left ribbon if and only if the framing goes, and stays, on the left-hand side.

6 Holonomy

According to [15, 16], the holonomy of a Hopf algebra gauge theory on Σ over the quasitriangular Hopf algebra Q is a groupoid homomorphism

$$\mathrm{Hol}\colon\ \mathsf{Path}(\Sigma) \to \mathrm{Hom}_K^\times\left(Q^*, \bigotimes_{e \in \Sigma^1} Q^*\right),$$

where the $\mathrm{Hom}_K(C, A)$ for a coalgebra C and algebra A is the convolution algebra [20] and $\mathrm{Hom}_K^\times(C, A)$ refers to the group of convolution invertible elements of $\mathrm{Hom}_K(C, A)$. For a gauge theory over $\mathcal{D}(H)$ we must choose $Q = \mathcal{D}(H)$. But since we need holonomy in the Kitaev model merely for the purpose of constructing string operators, we will follow the above recipe very loosely and use the name holonomy for any groupoid homomorphism

$$\mathrm{Hol}\colon\ \mathsf{Path}(\mathcal{D}(\Sigma)^*) \to \mathrm{Hom}_K^\times(\mathcal{D}(H)^*, \mathcal{M}(\Sigma)).$$

The convolution product of two functions $f, g\colon \mathcal{D}(H)^* \to \mathcal{M}$ is $f * g(\Psi) = f\left(\Psi_{(1)}\right)g\left(\Psi_{(2)}\right)$ for $\Psi \in \mathcal{D}(H)^*$ and the holonomy of a path $(d_n, \ldots, d_1)$ on $\mathcal{D}(\Sigma)^*$ is uniquely determined by the

holonomies of the length 1 paths by

$$\mathrm{Hol}_{(d_n,\dots,d_1)} = \mathrm{Hol}_{(d_n)} * \cdots * \mathrm{Hol}_{(d_2)} * \mathrm{Hol}_{(d_1)}. \tag{6.1}$$

In order for this formula to depend only on the path of the curve the length 1, holonomies should satisfy that $\mathrm{Hol}_{(\tilde{\mathbb{T}}_1 d)}$ be the convolution inverse of $\mathrm{Hol}_{(d)}$. If the function $\mathrm{Hol}_{(d)} \colon \mathcal{D}(H)^* \to \mathcal{M}$ happens to be either an algebra homomorphism or an antialgebra homomorphism, then its convolution inverse is simply $\mathrm{Hol}_{(d)} \circ S_{\mathcal{D}^*}$.

   We shall use also opholonomies

$$\mathrm{Hol}^\circ \colon\ \mathsf{Path}(\mathcal{D}(\Sigma)^*)^{\mathrm{op}} \to \mathrm{Hom}_K^\times(\mathcal{D}(H)^*, \mathcal{M}(\Sigma)),$$

which can be defined as functions on opcurves $(d_1, \dots, d_n)$ by

$$\mathrm{Hol}^\circ_{(d_1,\dots,d_n)} = \mathrm{Hol}^\circ_{(d_1)} * \mathrm{Hol}^\circ_{(d_2)} * \cdots * \mathrm{Hol}^\circ_{(d_n)}. \tag{6.2}$$

The condition of descending to the paths of opcurves is again that $\mathrm{Hol}^\circ_{(\tilde{\mathbb{T}}_1 d)}$ be the convolution inverse of $\mathrm{Hol}^\circ_{(d)}$.

   The ribbon operators of [1], as well as the symbolic arrows (3.31), suggest that opholonomy and holonomy should be defined by

$$\mathrm{Hol}^\circ_{(d)}(h \otimes \varphi) := \begin{cases} \varphi(1) P_a(h) & \text{if } d = a_0^+, \\ \varepsilon(h) Q_a(\varphi) & \text{if } d = a_2^+, \\ \varphi(1) P_a(S(h)) & \text{if } d = a_0^-, \\ \varepsilon(h) Q_a(S(\varphi)) & \text{if } d = a_2^- \end{cases} \tag{6.3}$$

for all $a \in \mathrm{Arr}(\Sigma)$ and

$$\mathrm{Hol}_{(d)} := S_{\mathcal{M}} \circ \mathrm{Hol}^\circ_{(d)} \circ S_{\mathcal{D}^*}, \qquad d \in \mathrm{Arr}(\mathcal{D}(\Sigma)^*), \tag{6.4}$$

where $S_{\mathcal{M}}$ is the involutive algebra automorphism $\mathcal{M} \to \mathcal{M}$ which sends $P_a(h)$ to $P_{T_1 a}(h)$ and $Q_a(\varphi)$ to $Q_{T_1 a}(\varphi)$. By (3.1), the $\mathrm{Hol}^\circ_{(d)}$ are antialgebra maps for $d = a_0^+$ and algebra maps for $d = a_2^+$. Moreover, in spite of the complicated expression (3.3) it is easy to see that $\mathrm{Hol}^\circ_{(\tilde{\mathbb{T}}_1 d)} = \mathrm{Hol}^\circ_{(d)} \circ S_{\mathcal{D}^*}$. So (6.3) together with (6.2) defines indeed an opholonomy. Then (6.4) together with (6.1) is automatically a holonomy.

   The relation (6.4) between holonomy and opholonomy is rather arbitrary. We could have defined them in such a way that $\mathrm{Hol}_{(d)} = \mathrm{Hol}^\circ_{(d)}$ for all arrows d of $\mathcal{D}(\Sigma)^*$. The choice (6.4) perhaps allows easier comparison with the existing definitions of the literature.

Definition 6.1. Let $\gamma = (d_1, \dots, d_n)$ be a (op)curve on $\mathcal{D}(\Sigma)^*$ with $d_k = (a_k)_{i_k}^{\sigma_k}$. Then γ is called proper if for all $1 \le k < l \le n$ the following statements hold: $a_k \ne a_l$ and if $a_k = T_1 a_l$ then $i_k = i_l$. In other words, no pairs d_k, d_l of arrows in the (op)curve are equal or incident and lie on the boundary of the same type 1 face $e \in \Sigma^1$.

Proposition 6.2. *The relation of proper (op)curves to ribbons is the following:*

(i) *Every proper curve $\gamma = (d_n, \dots, d_1)$ on $\mathcal{D}(\Sigma)^*$ is a ribbon curve and*

$$\mathrm{Hol}_\gamma = S_{\mathcal{M}} \circ \mathrm{Hol}^\circ_{\gamma^{\mathrm{rev}}} \circ S_{\mathcal{D}^*} = S_{\mathcal{M}} \circ \mathrm{Hol}^\circ_{\gamma^{\mathrm{rev}-1}}.$$

(ii) *A ribbon opcurve $\rho = (d_1, \dots, d_n)$ on $\mathcal{D}(\Sigma)^*$ is proper if and only if it is a simple opcurve. So proper ribbon and simple ribbon are synonymous.*

Proof. (i) Notice that an opposite pair d, $\tilde{\mathbb{T}}_1 d$ is always of the form a_i^+, a_i^-. Therefore, proper (op)curves are reduced. Hence, the property of being proper for consecutive arrows d_k, d_{k+1} is equivalent to the property of Lemma 5.8 (iii). This proves that proper (op)curves are ribbons. The property of being proper for any pair d_k, d_l implies that the P or Q operators associated to them by the opholonomy commute: $\left[\mathrm{Hol}^\circ_{(d_k)}(\Phi), \mathrm{Hol}^\circ_{(d_l)}(\Psi)\right] = 0$ for every $\Phi, \Psi \in \mathcal{D}(H)^*$ and for $k \neq l$. Therefore,

$$\mathrm{Hol}_\gamma(\Psi) = S_\mathcal{M}\big(\mathrm{Hol}^\circ_{(d_1)}\big(S_{\mathcal{D}^*}\big(\Psi_{(n)}\big)\big) \cdots \mathrm{Hol}^\circ_{(d_n)}\big(S_{\mathcal{D}^*}\big(\Psi_{(1)}\big)\big)\big) = S_\mathcal{M}\big(\mathrm{Hol}^\circ_{\gamma^{\mathrm{rev}}}(S_{\mathcal{D}^*}(\Psi))\big).$$

This proves the first equation. The second follows from the first by the observation that for a proper opcurve α the $\mathrm{Hol}^\circ_\alpha \circ S_{\mathcal{D}^*}$ is the convolution inverse of $\mathrm{Hol}^\circ_\alpha$ hence equal to the opholonomy of α^{-1}.

(ii) Assume ρ is not simple. Then there exists a site s and a minimal closed subcurve $(d_i, d_{i+1}, \ldots, d_{j-1})\colon s \to s$ of ρ. Here $i = 1$ or $j = n + 1$ is allowed but not both because this would imply that ρ is closed and simple. If $i > 1$, then d_{i-1} and d_{j-1} are incident to s and lie on the boundary of a type 1 face. If $j \leq n$, then the same holds for the arrows d_i and d_j. Anyway, we get contradiction with properness of ρ.

If ρ is not proper, then for some $1 \leq k < l \leq n$ the $d_k = (a_k)^{\sigma_k}_{i_k}$ and $d_l = (a_l)^{\sigma_l}_{i_l}$ satisfy one of the following: either $a_k = a_l$ or $(a_k = T_1 a_l$ and $i_k \neq i_l)$. In both cases ρ is either not simple or not a ribbon. ∎

En passant, we mention a useful rewriting of the exchange relations of the P and Q operators in terms of length 1 opholonomies:

$$\mathrm{Hol}^\circ_{(a_0^-)}(\Psi)\, \mathrm{Hol}^\circ_{(a_2^+)}(\Phi) = \mathrm{Hol}^\circ_{(a_2^+)}(\Phi \leftharpoonup R_1)\, \mathrm{Hol}^\circ_{(a_0^-)}(\Psi \leftharpoonup R_2), \tag{6.5}$$

$$\mathrm{Hol}^\circ_{((T_1 a)_2^+)}(\Psi)\, \mathrm{Hol}^\circ_{(a_0^-)}(\Phi) = \mathrm{Hol}^\circ_{(a_0^-)}(R_2 \rightharpoonup \Phi)\, \mathrm{Hol}^\circ_{((T_1 a)_2^+)}(R_1 \rightharpoonup \Psi), \tag{6.6}$$

where R is the R-matrix and $\Phi, \Psi \in \mathcal{D}(H)^*$.

The minimal closed ribbon (op)curves are the face loops that encircle a face of $\mathcal{D}(\Sigma)^*$. Since these loops depend on a base point, i.e., a site $s(a)$, we parametrize them with the arrow $a \in \mathrm{Arr}(\Sigma)$. The opcurve versions of them are

$$\text{type 0 loop around } \mathcal{O}_0(a)\colon \quad \alpha_a = \big((T_0 a)_0^+, (T_0^2 a)_0^+, \ldots, (T_0^m a)_0^+\big), \tag{6.7}$$

$$\text{type 1 loop around } \mathcal{O}_1(a)\colon \quad \beta_a = \big(a_0^+, a_2^+, (T_1 a)_0^+, (T_1 a)_2^+\big), \tag{6.8}$$

$$\text{type 2 loop around } \mathcal{O}_2(a)\colon \quad \gamma_a = \big(a_2^+, (T_2 a)_2^+, \ldots, (T_2^{n-1} a)_2^+\big), \tag{6.9}$$

where $m = |\mathcal{O}_0(a)|$ and $n = |\mathcal{O}_2(a)|$. The type 0 and type 2 loops are oriented counterclockwise and the type 1 loop clockwise, just like in Figure 4. The α_a and γ_a are proper ribbons but β_a is maximally improper. In order to facilitate the computation of the opholonomies of these opcurves we rewrite the formula (6.3) as follows:

$$\mathrm{Hol}^\circ_{(a_i^\sigma)} = \theta_a \circ \pi_i \circ S_{\mathcal{D}^*}^{(1-\sigma)/2},$$

where

$$\pi_0, \pi_2\colon \mathcal{D}(H)^* \to \mathcal{D}(H)^*, \qquad \pi_0(h \otimes \varphi) := h \otimes \varepsilon \cdot \varphi(1), \qquad \pi_2(h \otimes \varphi) := \varepsilon(h) \cdot 1 \otimes \varphi,$$

are algebra endomorphisms of the dual of the double and the

$$\theta_a\colon \ H \otimes H^* \to \mathcal{M}(\Sigma), \qquad \theta_a(h \otimes \varphi) := P_a(h) Q_a(\varphi)$$

are linear injections.

Lemma 6.3. *The maps θ_a, $a \in \mathrm{Arr}(\Sigma)$, satisfy the following properties:*

(i) $\theta_a = \mathrm{Hol}^\circ_{(a_0^+, a_2^+)}$.

(ii) $\theta_{T_1 a} = S_\mathcal{M} \circ \theta_a$.

(iii) $\theta_a(\Phi)\theta_b(\Psi) = \theta_b(\Psi)\theta_a(\Phi)$ *for all* $\Phi, \Psi \in \mathcal{D}(H)^*$ *if* $\mathcal{O}_1(a) \neq \mathcal{O}_1(b)$.

(iv) *The image of* θ_a *is the edge algebra* $\mathcal{M}_e \cong H \# H^*$ *on which* $S_\mathcal{M}$ *acts as the* $\mathcal{L}^{-1} \circ \mathcal{R} = \mathcal{R}^{-1} \circ \mathcal{L}$ *of* (3.12) *and* (3.13)*. The* θ_a *identifies this action of* $S_\mathcal{M}$ *with the action of the antipode of* $\mathcal{D}(H)^*$ *via*

$$S_\mathcal{M} \circ \theta_a = \theta_a \circ S_{\mathcal{D}^*}.$$

(v) *The* θ_a *becomes an antialgebra map*

$$\theta_a(\Phi)\theta_a(\Psi) = \theta_a(\Psi \underset{R}{\bullet} \Phi)$$

if we endow $\mathcal{D}(H)^*$*, as a monoid in the braided monoidal category of right* $\mathcal{D}(H)$*-modules, with the braided-opposite multiplication*

$$\Psi \underset{R}{\bullet} \Phi := (\Phi \leftharpoonup R_1)(\Psi \leftharpoonup R_2).$$

Proof. (i) Applying (6.3) and using the expression (3.2) for the coproduct, it is easy to see that the equation $\mathrm{Hol}^\circ_{(a_0^+, a_2^+)}(h \otimes \varphi) = P_a(h)Q_a(\varphi)$ holds.

(ii) is essentially the definition of $S_\mathcal{M}$.

(iii) holds because $\mathcal{M}$ is the tensor product of the edge algebras $\mathcal{M}_e$ for $e \in \Sigma^1$.

(iv) can be seen by comparing formulas (3.12), (3.13) with the antipode formula (3.3).

(v) The braiding on $(g \otimes \psi) \otimes (h \otimes \varphi) \in \mathcal{D}(H)^* \otimes \mathcal{D}(H)^*$ acts by

$$
\begin{aligned}
[(h \otimes \varphi) \leftharpoonup R_1] \otimes [(g \otimes \psi) \leftharpoonup R_2] &= \langle \varepsilon \otimes x_i, h_{(1)} \otimes \xi_j \varphi_{(1)} \xi_k \rangle \big(S^{-1}(x_k) h_{(2)} x_j \otimes \varphi_{(2)}\big) \\
&\quad \otimes \langle \xi_i \otimes 1, g_{(1)} \otimes \xi_l \psi_{(1)} \xi_m \rangle \big(S^{-1}(x_m) g_{(2)} x_l \otimes \psi_{(2)}\big) \\
&= \big(S^{-1}(x_k) h x_j \otimes \varphi_{(2)}\big) \otimes \langle \xi_j \varphi_{(1)} \xi_k, g_{(1)} \rangle \big(g_{(2)} \otimes \psi\big) \\
&= \big(S^{-1}(g_{(3)}) h g_{(1)} \otimes (\varphi \leftharpoonup g_{(2)})\big) \otimes \big(g_{(4)} \otimes \psi\big),
\end{aligned}
$$

therefore

$$(g \otimes \psi) \underset{R}{\bullet} (h \otimes \varphi) = g_{(4)} S^{-1}(g_{(3)}) h g_{(1)} \otimes (\varphi \leftharpoonup g_{(2)}) \psi = h\big(\varphi_{(1)} \rightharpoonup g\big) \otimes \varphi_{(2)} \psi$$

(even for non-involutive H-s) is precisely the multiplication of the smash product applied to $(h \otimes \varphi) \otimes (g \otimes \psi) \in (H \# H^*) \otimes (H \# H^*)$. ∎

Part (v) of the lemma immediately implies (cf. [15, Theorem 7.6]).

Corollary 6.4. *The algebra* $\mathcal{M}^{\mathrm{op}}$ *is isomorphic to the tensor product* $\bigotimes_{e \in \Sigma^1} \mathcal{D}(H)^*_R$*, where the algebra* $\mathcal{D}(H)^*_R$ *is* $\mathcal{D}(H)^*$ *endowed with the braided opposite multiplication.*

Proposition 6.5. *The opholonomy of the type 0 and 2 face loops* α_a *and* γ_a *of* (6.7) *and* (6.9) *essentially reproduce the Gauss' law and flux operators of* (3.21) *and* (3.22)*, respectively,*

$$\mathrm{Hol}^\circ_{\alpha_a}(h \otimes \varphi) = \varphi(1) \cdot G_a(h), \qquad \mathrm{Hol}^\circ_{\gamma_a}(h \otimes \varphi) = \varepsilon(h) \cdot F_a(\varphi).$$

The opholonomy of a type 1 loop is

$$\mathrm{Hol}^\circ_{\beta_a}(h \otimes \varphi) = \mathbf{1}_\mathcal{M} \cdot \langle R_1 S_\mathcal{D}(R_2), h \otimes \varphi \rangle, \tag{6.10}$$

where $R_1 S_\mathcal{D}(R_2) = S_\mathcal{D}(u)$ *with* u *denoting the Drinfeld element* $S_\mathcal{D}(R_2) R_1$ *of the quasitriangular Hopf algebra* $\langle \mathcal{D}(H), R \rangle$.

Proof. Iterating the formula $(\pi_0 \otimes \mathrm{id}) \circ \Delta_{\mathcal{D}^*}(h \otimes \varphi) = (h_{(1)} \otimes \varepsilon) \otimes (h_{(2)} \otimes \varphi)$ for multiple coproducts we arrive to

$$
\begin{aligned}
\mathrm{Hol}^\circ_{\alpha_a}(h \otimes \varphi) &= \theta_{T_0 a}(h_{(1)} \otimes \varepsilon) \cdots \theta_{T_0^m a}(h_{(m)} \otimes \varepsilon) \cdot \varphi(1) \\
&= P_{T_0 a}(h_{(1)}) \cdots P_{T_0^m a}(h_{(m)}) \cdot \varphi(1) = G_a(h) \cdot \varphi(1).
\end{aligned}
$$

Similarly, formula $(\mathrm{id} \otimes \pi_2) \circ \Delta_{\mathcal{D}^*}(h \otimes \varphi) = (h \otimes \varphi_{(1)}) \otimes (1 \otimes \varphi_{(2)})$ leads to

$$
\begin{aligned}
\mathrm{Hol}^\circ_{\gamma_a}(h \otimes \varphi) &= \varepsilon(h) \cdot \theta_a(1 \otimes \varphi_{(1)}) \cdots \theta_{T_2^{n-1} a}(1 \otimes \varphi_{(n)}) \\
&= \varepsilon(h) \cdot Q_a(\varphi_{(1)}) \cdots Q_{T_2^{n-1} a}(\varphi_{(n)}) = \varepsilon(h) \cdot F_a(\varphi).
\end{aligned}
$$

Using (i), (ii) and (iv) of Lemma 6.3, the opholonomy of β_a can be written as

$$
\mathrm{Hol}^\circ_{\beta_a} = \mathrm{Hol}^\circ_{(a_0^+, a_2^+)} * (S_{\mathcal{M}} \circ \mathrm{Hol}^\circ_{(a_0^+, a_2^+)}) = \theta_a * (\theta_a \circ S_{\mathcal{D}^*}),
$$

which, by (v) of the lemma, evaluates to

$$
\begin{aligned}
\mathrm{Hol}^\circ_{\beta_a}(h \otimes \varphi) &= \theta_a\big(\big((h \otimes \varphi)_{(1)} \leftharpoonup R_1\big)\big(S_{\mathcal{D}^*}\big((h \otimes \varphi)_{(2)}\big) \leftharpoonup R_2\big)\big) \\
&= \theta_a\big((h \otimes \varphi)_{(2)} S_{\mathcal{D}^*}\big((h \otimes \varphi)_{(3)}\big)\big)\big\langle (h \otimes \varphi)_{(1)}, R_1\big\rangle\big\langle S_{\mathcal{D}^*}\big((h \otimes \varphi)_{(4)}\big), R_2\big\rangle \\
&= \theta_a(1_{\mathcal{D}^*}) \cdot \langle h \otimes \varphi, R_1 S_{\mathcal{D}}(R_2)\rangle.
\end{aligned}
$$

We note that $R_1 S_{\mathcal{D}}(R_2) = (\varepsilon \otimes x_i)(S(\xi_i) \otimes 1) = (\varepsilon \otimes S(x_i))(\xi_i \otimes 1) = S_{\mathcal{D}}(R_1) R_2 = S_{\mathcal{D}}(u)$. ∎

Since the only non-commuting P and Q operators are lying at corners of type 1 loops, the result (6.10) can be thought of as a holonomy version of the Heisenberg exchange relations (3.7).

Using Proposition 6.2 (i), the above results can be translated to holonomies of the basic loop curves $\alpha_a^{\mathrm{rev}-1}$, $\beta_a^{\mathrm{rev}-1}$ and $\gamma_a^{\mathrm{rev}-1}$ as follows:

$$
\begin{aligned}
\mathrm{Hol}_{\alpha_a^{\mathrm{rev}-1}}(h \otimes \varepsilon) &= S_{\mathcal{M}}(G_a(h)) = P_{T_1 T_0 a}(h_{(1)}) P_{T_1 T_0^2 a}(h_{(2)}) \cdots P_{T_1 T_0^m a}(h_{(m)}), \\
\mathrm{Hol}_{\beta_a^{\mathrm{rev}-1}}(h \otimes \varphi) &= \mathbf{1}_{\mathcal{M}} \cdot \langle R_1 S_{\mathcal{D}}(R_2), h \otimes \varphi\rangle, \\
\mathrm{Hol}_{\gamma_a^{\mathrm{rev}-1}}(1 \otimes \varphi) &= S_{\mathcal{M}}(F_a(\varphi)) = Q_{T_1 a}(\varphi_{(1)}) Q_{T_1 T_2 a}(\varphi_{(2)}) \cdots Q_{T_1 T_2^{n-1} a}(\varphi_{(n)}).
\end{aligned}
$$

The curve $\underset{\sim}{\alpha}_a^{\mathrm{rev}-1}$, if viewed on Σ, consists of arrows pointing toward the vertex $v = \mathcal{O}_0(a)$ and, as a P-curve, winds around v in clockwise direction. The curve $\gamma_a^{\mathrm{rev}-1}$ consists of arrows on the boundary of the face $f = \mathcal{O}_2(a)$ directed in clockwise direction. These basic loops and their holonomies fit to the conventions used in the literature [3, 15]. Although the automorphism $S_{\mathcal{M}}$ provides a direct connection to our conventions, the simplicity of (3.21) and (3.22) can be an excuse for our preference of opholonomies.

The type 1 loops β_a and $\beta_a^{\mathrm{rev}-1}$ belong to the following type of (op)curves:

Definition 6.6. A closed (op)curve ζ on $\mathcal{D}(\Sigma)^*$ is called a central (op)curve if its (op)holonomy is of the form $(h \otimes \varphi) \mapsto \mathbf{1}_{\mathcal{M}} \cdot \langle Z, h \otimes \varphi\rangle$ for some central element Z of $\mathcal{D}(H)$.

Notice that by convolution invertibility of (op)holonomies the Z in the above definition is necessarily invertible. Due to Proposition 6.2 (i), a proper ribbon opcurve ζ is central if and only if the proper ribbon curve ζ^{rev} is central.

Lemma 6.7. *Let $s \in \mathcal{D}(\Sigma)^{*0}$ be a site.*

(i) *The central opcurves $\zeta \colon s \to s$ form a monoid under concatenation and their paths form a group.*

(ii) *If $\zeta\colon s \to s$ is a central opcurve and $\alpha\colon s \to s'$ is any opcurve, then $\zeta' := \alpha^{-1}\zeta\alpha\colon s' \to s'$ is a central opcurve, too. Moreover, $\mathrm{Hol}^{\circ}_{\zeta'} = \mathrm{Hol}^{\circ}_{\zeta}$.*

Proof. (i) If $\zeta_j\colon s \to s$ have $\mathrm{Hol}^{\circ}_{\zeta_j}(\Psi) = \mathbf{1}_{\mathcal{M}} \cdot \langle Z_j, \Psi \rangle$ for $j = 1, 2$, then $\mathrm{Hol}^{\circ}_{\zeta_1\zeta_2}(\Psi) = \mathbf{1}_{\mathcal{M}} \cdot \langle Z_1 Z_2, \Psi \rangle$. If ζ is central, then so is ζ^{-1}.

(ii) If we can prove the statement for α having length 1, then the statement will follow by induction on the length of α. Let $\alpha = (d)$. Then

$$\mathrm{Hol}^{\circ}_{\zeta'}(\Psi) = \mathrm{Hol}^{\circ}_{(d)}\big(S_{\mathcal{D}^*}\big(\Psi_{(1)}\big)\big)\,\mathrm{Hol}^{\circ}_{\zeta}\big(\Psi_{(2)}\big)\,\mathrm{Hol}^{\circ}_{(d)}\big(\Psi_{(3)}\big)$$
$$= \mathrm{Hol}^{\circ}_{(d)}\big(S_{\mathcal{D}^*}\big(\Psi_{(1)}\big)\big)\,\mathrm{Hol}^{\circ}_{(d)}\big(\Psi_{(3)}\big) \cdot \big\langle Z, \Psi_{(2)}\big\rangle.$$

Evaluating $S_{\mathcal{D}^*}\big(\Psi_{(1)}\big) \otimes \langle Z, \Psi_{(2)}\rangle \Psi_{(3)}$ on $Y \otimes Y' \in \mathcal{D}(H) \otimes \mathcal{D}(H)$, one obtains $\langle \Psi, S_{\mathcal{D}}(Y)ZY' \rangle = \langle \Psi, S_{\mathcal{D}}(Y)Y'Z \rangle$, therefore

$$S_{\mathcal{D}^*}\big(\Psi_{(1)}\big) \otimes \big\langle Z, \Psi_{(2)}\big\rangle \Psi_{(3)} = S_{\mathcal{D}^*}\big(\Psi_{(1)}\big) \otimes \Psi_{(2)}\big\langle Z, \Psi_{(3)}\big\rangle$$

and

$$\mathrm{Hol}^{\circ}_{\zeta'}(\Psi) = \mathrm{Hol}^{\circ}_{(d)}\big(S_{\mathcal{D}^*}\big(\Psi_{(1)}\big)\big)\,\mathrm{Hol}^{\circ}_{(d)}\big(\Psi_{(2)}\big) \cdot \big\langle Z, \Psi_{(3)}\big\rangle$$
$$= \mathbf{1}_{\mathcal{M}} \cdot \varepsilon_{\mathcal{D}^*}\big(\Psi_{(1)}\big)\big\langle Z, \Psi_{(2)}\big\rangle = \mathbf{1}_{\mathcal{M}} \cdot \big\langle Z, \Psi\big\rangle$$
$$= \mathrm{Hol}^{\circ}_{\zeta}(\Psi). \qquad\blacksquare$$

Definition 6.8. Let $\gamma\colon s \to s'$ be an opcurve on $\mathcal{D}(\Sigma)^*$. A central deformation of γ is an opcurve $\delta\colon s \to s'$ such that $\delta\gamma^{-1}$ is central.

7 The algebra of ribbon operators

In this section, we shall work with ribbon operators $M = \mathrm{Hol}^{\circ}_{\rho}(\Psi)$ for some ribbon path ρ and some $\Psi \in \mathcal{D}(H)^*$. Ribbon operators can be multiplied either using the convolution product or the multiplication of $\mathcal{M}$. The relations of the former type come from the definition of opholonomy,

$$\mathrm{Hol}^{\circ}_{\alpha} * \mathrm{Hol}^{\circ}_{\beta} = \mathrm{Hol}^{\circ}_{\alpha\,\beta}, \qquad \alpha, \beta \in \mathsf{Ribb}^{\circ}_{L} \quad \text{or} \quad \alpha, \beta \in \mathsf{Ribb}^{\circ}_{R}.$$

So we can, for example, change the starting point of the Gauss' law and flux operators by conjugation in the convolution sense,

$$G_{T_0 a}(h) = P_{T_0 a}(S\big(h_{(1)}\big))G_a\big(h_{(2)}\big)P_{T_0 a}\big(h_{(3)}\big), \tag{7.1}$$
$$F_{T_2 a}(\varphi) = Q_a(S\big(\varphi_{(1)}\big))F_a\big(\varphi_{(2)}\big)Q_a\big(\varphi_{(3)}\big). \tag{7.2}$$

As for the ordinary multiplication of $\mathcal{M}$, it is quite difficult to find computable relations between ribbon operators. The rare cases which we are able to compute occupy the present Section.

If two ribbons are sufficiently distant, then the corresponding operators should commute. Thinking over which edges have P or Q operators that do not commute we arrive to the following:

Definition 7.1. For an arrow $d \in \mathrm{Arr}(\mathcal{D}(\Sigma)^*)$, let $U(d)$ be the set of arrows that are incident to d and lie on the boundary of the same type 1 face as d does. (There are 6 such arrows, including d, forming a U-shape of 3 edges.) Two (op)curves, or just sets of arrows, γ and γ' in $\mathcal{D}(\Sigma)^*$ are called U-separated, written $\gamma \underset{U}{\leftrightarrow} \gamma'$ if $d \notin U(d')$ for all $d \in \gamma$ and $d' \in \gamma'$.

So we have

$$\mathrm{Hol}^{\circ}_{\rho}(\Phi)\,\mathrm{Hol}^{\circ}_{\gamma}(\Psi) = \mathrm{Hol}^{\circ}_{\gamma}(\Psi)\,\mathrm{Hol}^{\circ}_{\rho}(\Phi) \qquad \forall \Phi, \Psi \in \mathcal{D}(H)^* \quad \text{if } \rho \underset{\mathrm{U}}{\leftrightarrow} \gamma. \tag{7.3}$$

At the other extreme let $\rho = \gamma$. Then the product is computable if $d \notin U(d')$ for any pair of different arrows $d, d' \in \rho$. This property of being "self-U-separated" can be seen to be equivalent to properness (Definition 6.1) of the opcurve. Thus,

$$\mathrm{Hol}^{\circ}_{\rho}(\Phi)\,\mathrm{Hol}^{\circ}_{\rho}(\Psi) = \mathrm{Hol}^{\circ}_{\rho}(\Phi\Psi) \qquad \text{if } \rho \in \mathsf{Ribb}^{\circ}_{L} \text{ is proper} \tag{7.4}$$

$$= \mathrm{Hol}^{\circ}_{\rho}(\Psi\Phi) \qquad \text{if } \rho \in \mathsf{Ribb}^{\circ}_{R} \text{ is proper}$$

by comparing (6.3) with (3.1) and using that the coproduct of $\mathcal{D}(H)^*$ is multiplicative. For proper ribbons, we also have the inversion formula

$$\mathrm{Hol}^{\circ}_{\rho^{-1}}(\Phi) = \mathrm{Hol}^{\circ}_{\rho}(S_{\mathcal{D}^*}(\Phi)) \qquad \text{if } \rho \text{ is proper} \tag{7.5}$$

and the cyclic permutation formula

$$\mathrm{Hol}^{\circ}_{C\rho}(\Psi) = \mathrm{Hol}^{\circ}_{\rho}(\Psi) \qquad \text{if } \rho \text{ is closed, proper and } \Psi \in \mathrm{Cocom}\,\mathcal{D}(H)^*, \tag{7.6}$$

where $C\rho = (d_2, \ldots, d_n, d_1)$ whenever $\rho = (d_1, \ldots, d_n)$.

Now consider the length 2 opcurves that are not straight, the "elbows". Up to taking inverses, there are 4 such opcurves with common middle point being the site $s(a)$:

$$\left(a_0^+, a_2^+\right), \qquad \left((T_2^{-1}a)_2^+, a_2^+\right), \qquad \left((T_2^{-1}a)_2^+, (T_0a)_0^+\right), \qquad \left(a_0^+, (T_0a)_0^+\right).$$

The following lemma is known in the context of Hopf algebra gauge theory [15].

Lemma 7.2. *The opholonomies of an "opposite pair of elbows" commute. That is to say,*

$$\left[\mathrm{Hol}^{\circ}_{(a_0^+, a_2^+)}(\Phi), \mathrm{Hol}^{\circ}_{((T_2^{-1}a)_2^+, (T_0a)_0^+)}(\Psi)\right] = 0, \tag{7.7}$$

$$\left[\mathrm{Hol}^{\circ}_{((T_2^{-1}a)_2^+, a_2^+)}(\Phi), \mathrm{Hol}^{\circ}_{(a_0^+, (T_0a)_0^+)}(\Psi)\right] = 0 \tag{7.8}$$

for all $\Phi, \Psi \in \mathcal{D}(H)^*$. If we replace any one or both of the elbows in (7.7) or (7.8) by their inverses commutativity of their opholonomies remains true.

Proof. (7.7) follows from (7.3) since the two curves are U-separated. (7.8) can be shown by checking that

$$\mathrm{Hol}^{\circ}_{((T_2^{-1}a)_2^+, a_2^+)}(g \otimes \varphi) = \varepsilon(g) \cdot Q_{T_2^{-1}a}(\varphi_{(1)}) Q_a(\varphi_{(2)}),$$

$$\mathrm{Hol}^{\circ}_{(a_0^+, (T_0a)_0^+)}(h \otimes \psi) = \psi(1) \cdot P_a(h_{(1)}) P_{T_0a}(h_{(2)})$$

and then computing

$$P_a(h_{(1)}) P_{T_0a}(h_{(2)}) \cdot Q_{T_2^{-1}a}(\varphi_{(1)}) Q_a(\varphi_{(2)}) =$$

$$\overset{(3.18)}{=} P_a(h_{(1)}) Q_{T_2^{-1}a}(\varphi_{(1)}) P_{T_0a}(h_{(2)} \leftharpoonup \varphi_{(2)}) Q_a(\varphi_{(3)})$$

$$= Q_{T_2^{-1}a}(\varphi_{(1)}) P_a(h_{(1)}) Q_a(\varphi_{(3)}) P_{T_0a}(h_{(2)} \leftharpoonup \varphi_{(2)})$$

$$\overset{(3.16)}{=} Q_{T_2^{-1}a}(\varphi_{(1)}) Q_a(\varphi_{(4)}) P_a(S(\varphi_{(3)}) \rightharpoonup h_{(1)}) P_{T_0a}(h_{(2)} \leftharpoonup \varphi_{(2)})$$

$$= Q_{T_2^{-1}a}(\varphi_{(1)}) Q_a(\varphi_{(2)}) \cdot P_a(h_{(1)}) P_{T_0a}(h_{(2)}).$$

Since the elbows in (7.8) are proper ribbons, replacing any one of them with its inverse amounts only to applying $S_{\mathcal{D}^*}$ on Φ or Ψ, see (7.5). Taking inverses of the elbows of (7.7) requires more work,

$$\mathrm{Hol}^{\circ}_{(a_0^+, a_2^+)^{-1}} = \mathrm{Hol}^{\circ}_{((T_1 a)_0^+, (T_1 a)_2^+)} * \mathrm{Hol}^{\circ}_{\beta_a^{-1}},$$

$$\mathrm{Hol}^{\circ}_{((T_2^{-1} a)_2^+, (T_0 a)_0^+)^{-1}} = \mathrm{Hol}^{\circ}_{((T_1 T_2^{-1} a)_2^+, (T_1 T_0 a)_0^+)} * \mathrm{Hol}^{\circ}_{\beta'_b^{-1}},$$

where $b = T_2^{-1} a$ and $\beta'_b := (b_0^-)\beta_b(b_0^+)$. The appearance of T_1 in the first terms does not change U-separatedness of the curves and the β_a and β'_b being central the second terms do not change commutativity, either. ∎

Corollary 7.3. *Let λ be either a type 0 or a type 2 face loop. Let $\epsilon = (d, b)$ be an elbow opcurve with midpoint $\nabla_1 d = \nabla_0 b = s$ which intersects λ only at the site s. If s is not the base point of λ, then $\mathrm{Hol}^{\circ}_{\epsilon}(\Phi)$ commutes with $\mathrm{Hol}^{\circ}_{\lambda}(\Psi)$ for every $\Phi, \Psi \in \mathcal{D}(H)^*$.*

The above corollary is the first instance of a more general commutativity theorem of holonomies of type 0 or type 2 face loops with certain ribbon operators; expressing a sort of "gauge invariance" of the ribbon operator. Before formulating the statement, let us recall some Hopf algebra.

If $h \in \mathrm{Cocom}\, H$, then the iterated coproducts $h_{(1)} \otimes \cdots \otimes h_{(n)}$ are invariant under cyclic permutations. In particular, the Gauss' law operators $G_a(h)$ are independent of the choice of the base point, $G_{T_0 a}(h) = G_a(h)$, if $h \in \mathrm{Cocom}\, H$. The same holds for the flux operators $F_a(\varphi)$ if $\varphi \in \mathrm{Cocom}\, H^*$. A cocommutative element φ defines a trace on H by the evaluation, $\langle \varphi, hg \rangle = \langle \varphi, gh \rangle$. All traces on H are of this form. Since the Haar integral ι is cocommutative and the map $H \to H^*$, $h \mapsto \iota \leftharpoonup h$ is an isomorphism [20], every $\varphi \in \mathrm{Cocom}\, H^*$ is of the form $\varphi = \iota \leftharpoonup z$ for a unique central element z of H.

The relation of cocommutative elements of H and H^* with that of $\mathcal{D}(H)^*$ is intricate. Suppose we want the holonomy of the type 0 face loop α_a to be independent of the base point. For which elements $\Psi \in \mathcal{D}(H)^*$ do we have the relation $\mathrm{Hol}^{\circ}_{\alpha_{T_0 a}}(\Psi) = \mathrm{Hol}^{\circ}_{\alpha_a}(\Psi)$? One answer is provided by (7.6) saying that $\Psi \in \mathrm{Cocom}\, \mathcal{D}(H)^*$. Another answer is obtained by Proposition 6.5 and by equation (7.1) saying that $\Psi = h \otimes \varepsilon$ with $h \in \mathrm{Cocom}\, H$. Which answer implies the other? None of them. The point is that the surjective coalgebra maps

$$H \xleftarrow{\quad \mathrm{id} \otimes \varepsilon \quad} \mathcal{D}(H)^* \xrightarrow{\quad \varepsilon \otimes \mathrm{id} \quad} H^*$$

if restricted to $\mathrm{Cocom}\, \mathcal{D}(H)^*$ become mere linear maps

$$\mathrm{Cocom}\, H \xleftarrow{\quad \mathrm{id} \otimes \varepsilon \quad} \mathrm{Cocom}\, \mathcal{D}(H)^* \xrightarrow{\quad \varepsilon \otimes \mathrm{id} \quad} \mathrm{Cocom}\, H^*,$$

which are no longer surjective, unless $\mathrm{Cocom}\, H$, resp. $\mathrm{Cocom}\, H^*$, is commutative. This is why condition A of the theorem below contains a list of 3 possibilities.

Theorem 7.4. *Let $\rho \colon s_0 \to s_1$ be a ribbon opcurve (Definition 5.9) and let λ be either a type 0 face loop or a type 2 face loop which is disjoint from both s_0 and s_1. Then*

$$\mathrm{Hol}^{\circ}_{\rho}(\Phi) \text{ and } \mathrm{Hol}^{\circ}_{\lambda}(\Psi) \text{ commute for all } \Phi \in \mathcal{D}(H)^*$$

provided one of the following conditions holds:

A. *Ψ belongs to either $\mathrm{Cocom}\, \mathcal{D}(H)^*$ or $\mathrm{Cocom}\, H \otimes \varepsilon$ or $1 \otimes \mathrm{Cocom}\, H^*$.*

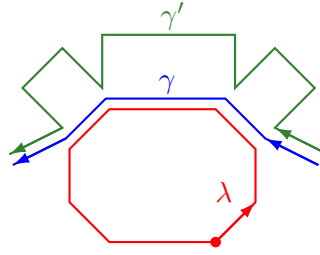B. *$\Psi \in \mathcal{D}(H)^*$ is arbitrary but the base point of λ is not lying on ρ.*

Figure 5. To the proof of Theorem 7.4: A subcurve γ of a left ribbon which collars the face loop λ of type 2 and its central deformation γ'.

Proof. First we assume that ρ is a left ribbon and λ is a type 2 face loop. Let $w_\rho \in \left\{ T_0^{-1}, T_2 \right\}^*$ be the code word of ρ (Definition 5.7). Let $\alpha_1, \ldots, \alpha_k$ be the list of generalized subcurves of ρ such that (1) all arrows and sites of α_j belongs to λ and (2) α_j is maximal among the subcurves satisfying (1). Mentioning sites under (1) refers to the possibility that α_j consists of a single site. It is maximal if and only if none of its neighbour arrows in ρ belongs to λ. The arrow d_j^{in} of ρ preceding α_j and the arrow d_j^{out} succeeding it must belong to $U(\lambda)$ but not to λ. These arrows are uniquely determined by the arc α_j as follows. If $\alpha_j = (s)$ is trivial, then $\left(d_j^{\mathrm{in}}, d_j^{\mathrm{out}} \right)$ is the unique elbow which touches λ only at the site s. If $\alpha_j = \left(c_{j,1}, \ldots, c_{j,l_j} \right)$ is non-trivial, then the $\left(d_j^{\mathrm{in}}, c_{j,1} \right)$ and $\left(c_{j,l_j}, d_j^{\mathrm{out}} \right)$ are straight (do not turn either left or right). The existence of such $d_j^{\mathrm{in/out}}$ in ρ for all j follows either from maximality of the arcs α_j or $\left(\text{for } d_1^{\mathrm{in}} \text{ and } d_k^{\mathrm{out}} \right)$ from the assumption that s_0, s_1 do not lie on λ. Let $\gamma_j := \left(d_j^{\mathrm{in}} \right) \alpha_j \left(d_j^{\mathrm{out}} \right)$. Then every γ_j corresponds to a subword $T_0^{-1} T_2^{l_j} T_0^{-1}$ of w_ρ and the whole ribbon can be written as the composite

$$\rho = \delta_0 \gamma_1 \delta_1 \cdots \gamma_k \delta_k$$

of possibly empty[1] curves δ_j U-separated from λ and of "collars" γ_j of λ. These collars need not be as simple as the one of Figure 5. They can wind around λ arbitrary many times, although always counterclockwise. For different indices, γ_i and γ_j may overlap arbitrarily or can even coincide.

It is now clear that commutativity of the opholonomies of ρ and λ will follow from that of γ and λ for a single collar $\gamma = \left(d^{\mathrm{in}}, c_1, \ldots, c_l, d^{\mathrm{out}} \right)$. A central deformation γ' of γ can be defined in the code word representation (Lemma 5.7) by means of a homomorphism $\mathfrak{h} \colon \left\{ T_0^{-1}, T_2 \right\}^* \to \left\{ T_0, T_0^{-1}, T_2, T_2^{-1} \right\}^*$ of free monoids as

$$\gamma' := \langle a_0, \mathfrak{h}(w_\gamma) \rangle, \qquad \text{where} \quad \mathfrak{h}\left(T_0^{-1} \right) := T_0^{-1}, \ \mathfrak{h}(T_2) := T_0^{-1} T_2^{-1} T_0^{-1},$$

where $\gamma = \langle a_0, w_\gamma \rangle$. This means that γ' differs from γ by replacing each c_j of color T_2 by a detour $\left(b_j, c_j', d_j \right)$ of color $\mathfrak{h}(T_2)$ such that the difference $T_2 \mathfrak{h}(T_2)^{-1} = T_2 T_0 T_2 T_0$ is a type 1 loop, a conjugate of some β_a, hence a central curve. To see that γ' is a central deformation of γ introduce the opcurves $\sigma_j := \left(d^{\mathrm{in}}, c_1, \ldots, c_j \right)$ and the lassos $\vartheta_j := \sigma_j \beta_j \sigma_j^{-1}$ with the central curve $\beta_j = \left(\tilde{\mathbb{T}}_1 c_j, b_j, c_j', d_j \right)^{-1}$. Then we have equality of paths

$$\gamma {\gamma'}^{-1} = \vartheta_l \cdots \vartheta_1$$

proving that $\gamma {\gamma'}^{-1}$ is a central opcurve by Lemma 6.7. Thus,

$$\mathrm{Hol}_\gamma^\circ = \mathrm{Hol}_{\gamma {\gamma'}^{-1}}^\circ * \mathrm{Hol}_{\gamma'}^\circ = \mathbf{1}_{\mathcal{M}} \cdot \langle Z, \_ \rangle * \mathrm{Hol}_{\gamma'}^\circ$$

---

[1]δ_j is empty if $\left(d_j^{\mathrm{out}}, d_{j+1}^{\mathrm{in}} \right)$ is a type 0 loop of length 2. Such double edges in $\mathcal{D}(\Sigma)^*$ correspond to a 2-valent vertex in Σ.

(for some $Z \in \mathrm{Center}(\mathcal{D}(H))$) commutes with $\mathrm{Hol}^\circ_\lambda(\Psi)$ if and only if $\mathrm{Hol}^\circ_{\gamma'}$ does. We write γ' as the composite

$$(d_0, b_1)(c'_1)(d_1, b_2)(c'_2)\cdots(c'_l)(d_l, b_{l+1}), \qquad \text{where} \quad d_0 = d^{\mathrm{in}}, \quad b_{l+1} = d^{\mathrm{out}}$$

of elbows (d_j, b_{j+1}) touching λ at a single site and of length 1 pieces (c'_j). The elbows commute with $\mathrm{Hol}^\circ_\lambda(\Psi)$ because – in case A – the assumption on Ψ ensures $\mathrm{Hol}^\circ_\lambda(\Psi)$ is independent of the base point of λ so that it can be shifted to become different from the contact point of the elbow and Corollary 7.3 to apply – and in case B – the base point of λ is different from the contact point and Corollary 7.3 directly applies. We claim that (c'_j) is U-separated from λ. Assume it is not. Then at least one endpoint of c'_j lies on λ so either b_j or d_j connects 2 sites of λ. Let $s(a_1)$ and $s(a_2)$ be these two sites. Since they lie on the face loop λ, a_1 and a_2 must belong to the same T_2-orbit on $\mathrm{Arr}(\Sigma)$. But $s(a_1)$ and $s(a_2)$ are also connected by b_j or d_j which have color T_0^{-1}, so a_1 and a_2 belong to the same T_0-orbit. By Lemma 2.2 (iii), this is possible only if $a_1 = a_2$ so b_j or d_j is a loop which cannot exist on $\mathcal{D}(\Sigma)^*$. This contradiction refutes the assumption that (c'_j) is not U-separated from λ. Hence, $\mathrm{Hol}^\circ_\lambda(\Psi)$ commutes also with the $\mathrm{Hol}^\circ_{(c'_j)}(\Phi)$ for all j and finally also with $\mathrm{Hol}^\circ_{\gamma'}$. This finishes the proof of the theorem for left ribbons ρ and type 2 face loops λ.

If ρ is a left ribbon and λ is a type 0 face loop, then the proof goes exactly as above except that the collars γ now, having color code $T_2(T_0^{-1})^l T_2$, wind around λ in clockwise direction. The deformation γ' is given by the substitution $T_0^{-1} \mapsto T_2 T_0 T_2$, $T_2 \mapsto T_2$.

If ρ is a right ribbon, then the collars go clockwise around type 2 face loops and counterclockwise around type 0 loops. Otherwise, the proof is the same. ∎

Corollary 7.5. *For every ribbon $\rho\colon \langle v_0, f_0 \rangle \to \langle v_1, f_1 \rangle$ and for every $\Phi \in \mathcal{D}(H)^*$, the operator $M = \mathrm{Hol}^\circ_\rho(\Phi)$ commutes with all the operators A_v and B_f except for $v \in \{v_0, v_1\}$ and $f \in \{f_0, f_1\}$.*

For a given ribbon, there are four type 0 or type 2 face loops for which the above theorem does not apply. The next theorem is to deal with these remaining cases.

Theorem 7.6. *Let $\rho\colon s(a) \to s(b)$ be a proper left ribbon, where $s(a)$ and $s(b)$ are disjoint sites. Then*

$$\mathrm{Hol}^\circ_\rho(\Phi) \underset{a}{\triangleleft} X = \mathrm{Hol}^\circ_\rho(\Phi \leftharpoonup X), \tag{7.9}$$

$$X \underset{b}{\triangleright} \mathrm{Hol}^\circ_\rho(\Phi) = \mathrm{Hol}^\circ_\rho(X \rightharpoonup \Phi) \tag{7.10}$$

hold true for every $X \in \mathcal{D}(H)$ and $\Phi \in \mathcal{D}(H)^*$.

Proof. The equations to be proven are equivalent to the exchange relations

$$\mathrm{Hol}^\circ_\rho(\Phi)\mathbf{D}_a(X) = \mathbf{D}_a(X_{(1)})\,\mathrm{Hol}^\circ_\rho(\Phi \leftharpoonup X_{(2)}), \tag{7.11}$$

$$\mathbf{D}_b(X)\,\mathrm{Hol}^\circ_\rho(\Phi) = \mathrm{Hol}^\circ_\rho(X_{(1)} \rightharpoonup \Phi)\mathbf{D}_b(X_{(2)}). \tag{7.12}$$

Every arrow of ρ is either a c_0^- or a c_2^+ for some $c \in \mathrm{Arr}(\Sigma)$. The consecutive c_0^- arrows form a non-trivial arc on a type 0 face loop and the consecutive c_2^+ arrows do the same on a type 2 face loop. So there is a unique decomposition $\rho = v_1 \cdots v_m$ into a concatenation of arcs which are alternatingly type 0 or type 2. (This decomposition corresponds to the decomposition of a word $w \in \{T_0^{-1}, T_2\}^*$ into an alternating product of powers T_0^{-k} and T_2^l.) Since ρ is proper, two arcs can intersect only at some of their common endpoints and in this case the two arcs are consecutive in the sequence $(v_1, \ldots, v_m)$. The number of arcs $m \geq 2$ because we assumed that $s(a)$ and $s(b)$ are disjoint.

The face loops α_a, γ_a and α_b, γ_b involved in the gauge transformations $\mathbf{D}_a(X)$ and $\mathbf{D}_b(X)$ fail to commute only with a few initial and final arcs. More precisely, let $v_j\colon s_{j-1} \to s_j$ and assume that v_1 is type 2. Then s_1 does not lie on α_a, otherwise two different sites, s_0 and s_1, belonged to both a type 0 and a type 2 face loop which is impossible in view of Figure 4. So neither endpoints of $v_2 \cdots v_m$ belongs to α_a. Furthermore, $v_2 \cdots v_m$ never visits the site $s(a) = s_0$ because v_1 does but ρ is proper, hence a simple curve by Proposition 6.2 (ii). It follows that

$$v_1 \text{ is type } 2 \quad \Rightarrow \quad G_a(H) \text{ commutes with } \mathrm{Hol}^{\circ}_{v_2\dots v_m}(\mathcal{D}(H)^*) \tag{7.13}$$

by Theorem 7.4 (B). Now assume that v_1 is type 0. Then s_1 lies on α_a and s_2 does not because v_2 is type 2. Therefore,

$$v_1 \text{ is type } 0 \quad \Rightarrow \quad G_a(H) \text{ commutes with } \mathrm{Hol}^{\circ}_{v_3\dots v_m}(\mathcal{D}(H)^*) \tag{7.14}$$

again by Theorem 7.4 (B). Similar arguments yield

$$v_1 \text{ is type } 0 \quad \Rightarrow \quad F_a(H^*) \text{ commutes with } \mathrm{Hol}^{\circ}_{v_2\dots v_m}(\mathcal{D}(H)^*), \tag{7.15}$$
$$v_1 \text{ is type } 2 \quad \Rightarrow \quad F_a(H^*) \text{ commutes with } \mathrm{Hol}^{\circ}_{v_3\dots v_m}(\mathcal{D}(H)^*). \tag{7.16}$$

Similar conclusions can be drawn for commutativity of $F_b(H^*)$ and $G_b(H)$ with $\mathrm{Hol}^{\circ}_{v_1\dots v_{m-1}}$ or $\mathrm{Hol}^{\circ}_{v_1\dots v_{m-2}}$ but we shall not need them.

Concentrating on the relation (7.11) we have 6 different exchange relations to compute:

Case G_a with v_1 if v_1 is type 0. Let $n = |\mathcal{O}_0(a)|$. Then the arc is $v_1 = \big(a_0^-, (T_0^{-1}a)_0^-, \dots, (T_0^{-j+1}a)_0^-\big)$ for some $0 < j < n$. Since $\mathrm{Hol}^{\circ}_{v_1}(h \otimes \varphi)$ is proportional to $\varphi(1)$, it suffices to work with $\mathrm{Hol}^{\circ}_{v_1}(h\otimes\varepsilon) = P^S_{a,\dots,T_0^{-j+1}a}(h)$ where we use the notation $P_{c_1,\dots,c_i}(h) := P_{c_1}\big(h_{(1)}\big)\cdots P_{c_i}\big(h_{(i)}\big)$ and similarly for $P^S_c = P_c \circ S$,

$$\begin{aligned}
P^S_{a,\dots,T_0^{-j+1}a}(h)G_a(g) &= P_{T_0^{n-j+1}a,\dots,T_0^n a}(S(h))P_{T_0 a,\dots,T_0^{n-j}a}\big(g_{(1)}\big)P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(g_{(2)}\big) \\
&= P_{T_0 a,\dots,T_0^{n-j}a}\big(g_{(1)}\big)P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(S(h)g_{(2)}\big) \\
&= G_a\big(g_{(1)}\big)P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(S\big(g_{(2)}\big)S(h)g_{(3)}\big) \\
&= G_a\big(g_{(1)}\big)P^S_{a,\dots,T_0^{-j+1}a}\big(S\big(g_{(3)}\big)hg_{(2)}\big).
\end{aligned}$$

Case G_a with v_2 if v_1 is type 0. Given v_1 as in the previous case the next arc v_2 starts at $s\big(T_0^{-j}a\big)$ therefore $v_2 = \big((T_0^{-j}a)_2^+, (T_2 T_0^{-j}a)_2^+, \dots, (T_2^{k-1}T_0^{-j}a)_2^+\big)$ for some $0 < k < \big|\mathcal{O}_2(T_0^{-j}a)\big|$. Nontriviality of the commutator of $\mathrm{Hol}^{\circ}_{v_2}(1 \otimes \varphi) = Q_{T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}(\varphi)$ and G_a comes from the first arrow of v_2 and from the $(n-j)$-th arrow of G_a,

$$\begin{aligned}
Q_{T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}(\varphi)G_a(g) &= Q_{T_0^{-j}a}\big(\varphi_{(1)}\big)G_a(g)Q_{T_2 T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}\big(\varphi_{(2)}\big) \\
&= P_{T_0 a,\dots,T_0^{n-j-1}a}\big(g_{(1)}\big)Q_{T_0^{-j}a}\big(\varphi_{(1)}\big)P_{T_0^{n-j}a}\big(g_{(2)}\big) \\
&\quad \times P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(g_{(3)}\big)Q_{T_2 T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}\big(\varphi_{(2)}\big) \\
&\overset{(2.8)}{=} P_{T_0 a,\dots,T_0^{n-j-1}a}\big(g_{(1)}\big)P_{T_0^{n-j}a}\big(\varphi_{(1)} \rightharpoonup g_{(2)}\big)Q_{T_0^{-j}a}\big(\varphi_{(2)}\big) \\
&\quad \times P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(g_{(3)}\big)Q_{T_2 T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}\big(\varphi_{(3)}\big) \\
&= P_{T_0 a,\dots,T_0^{n-j}a}\big(g_{(1)}\big)P_{T_0^{n-j+1}a,\dots,T_0^n a}\big(g_{(3)}\big)Q_{T_0^{-j}a,\dots,T_2^{k-1}T_0^{-j}a}\big(\varphi \leftharpoonup g_{(2)}\big).
\end{aligned}$$

This is not yet in the desired form, there is no G_a on the left-hand side, but we wish to complete it in the presence of the holonomy of v_1:

$$
\begin{aligned}
&P^S_{a,\ldots,T_0^{-j+1}a}(h)Q_{T_0^{-j}a,\ldots,T_2^{k-1}T_0^{-j}a}(\varphi)G_a(g)\\
&= P_{T_0^{n-j+1}a,\ldots,T_0^n a}(S(h))P_{T_0 a,\ldots,T_0^{n-j}a}(g_{(1)})P_{T_0^{n-j+1}a,\ldots,T_0^n a}(g_{(3)})Q_{T_0^{-j}a,\ldots,T_2^{k-1}T_0^{-j}a}(\varphi \leftharpoonup g_{(2)})\\
&= P_{T_0 a,\ldots,T_0^{n-j}a}(g_{(1)})P_{T_0^{n-j+1}a,\ldots,T_0^n a}(S(h)g_{(3)})Q_{T_0^{-j}a,\ldots,T_2^{k-1}T_0^{-j}a}(\varphi \leftharpoonup g_{(2)})\\
&= G_a(g_{(1)})P_{T_0^{n-j+1}a,\ldots,T_0^n a}(S(g_{(2)})S(h)g_{(4)})Q_{T_0^{-j}a,\ldots,T_2^{k-1}T_0^{-j}a}(\varphi \leftharpoonup g_{(3)})\\
&= G_a(g_{(1)})P^S_{a,\ldots,T_0^{-j+1}a}(S(g_{(4)})hg_{(2)})Q_{T_0^{-j}a,\ldots,T_2^{k-1}T_0^{-j}a}(\varphi \leftharpoonup g_{(3)}).
\end{aligned}
$$

Since $\mathrm{Hol}^\circ_{v_1 v_2}(\Phi) = \mathrm{Hol}^\circ_{v_1}(h \otimes \varepsilon)\,\mathrm{Hol}^\circ_{v_2}(1 \otimes \varphi)$, where $\Phi = h \otimes \varphi$ and

$$
S(g_{(3)})hg_{(1)} \otimes \varphi \leftharpoonup g_{(2)} = \langle h_{(1)} \otimes \xi_i \varphi_{(1)} \xi_j, \varepsilon \otimes g \rangle S(x_j)h_{(2)}x_i \otimes \varphi_{(2)} = \Phi \leftharpoonup (\varepsilon \otimes g),
$$

we obtain $\mathrm{Hol}^\circ_{v_1 v_2}(\Phi)G_a(g) = G_a(g_{(1)})\,\mathrm{Hol}^\circ_{v_1 v_2}(\Phi \leftharpoonup (\varepsilon \otimes g_{(2)}))$ which, together with (7.14), gives the result

$$
\mathrm{Hol}^\circ_\rho(\Phi)G_a(g) = G_a(g_{(1)})\,\mathrm{Hol}^\circ_\rho(\Phi \leftharpoonup (\varepsilon \otimes g_{(2)})), \tag{7.17}
$$

whenever v_1 is a type 0 arc.

Case F_a with v_1 if v_1 is type 0. For v_1 as above, we can write

$$
\begin{aligned}
P^S_{a,\ldots,T_0^{-j+1}a}(h)F_a(\psi) &= P_a(S(h_{(1)}))F_a(\psi)P_{T_0^{-1}a}(S(h_{(2)})) \cdots P_{T_0^{-j+1}a}(S(h_{(j)}))\\
&\overset{(3.13)}{=} F_a(\psi_{(2)})P_a(S(\psi_{(1)}) \rightharpoonup S(h_{(1)}))P_{T_0^{-1}a}(S(h_{(2)})) \cdots P_{T_0^{-j+1}a}(S(h_{(j)}))\\
&= F_a(\psi_{(2)})P^S_{a,\ldots,T_0^{-j+1}a}(h \leftharpoonup \psi_{(1)}).
\end{aligned}
$$

Since

$$
\begin{aligned}
\Phi \leftharpoonup (\psi \otimes 1) &= \langle \psi, h_{(1)} \rangle h_{(2)} \otimes \varphi = h \leftharpoonup \psi \otimes \varphi,\\
\mathrm{Hol}^\circ_{v_1}(\Phi \leftharpoonup (\psi \otimes 1)) &= P^S_{a,\ldots,T_0^{-j+1}a}(h \leftharpoonup \psi) \cdot \langle \varphi, 1 \rangle,
\end{aligned}
$$

using also (7.15), we immediately obtain

$$
\mathrm{Hol}^\circ_\rho(\Phi)F_a(\psi) = F_a(\psi_{(2)})\,\mathrm{Hol}^\circ_\rho(\Phi \leftharpoonup (\psi_{(1)} \otimes 1)), \tag{7.18}
$$

which together with (7.17) proves (7.11) at least if the first arc of ρ is type 0.

Case G_a with v_1 if v_1 is type 2. Then $v_1 = (a_2^+, \ldots, (T_2^{j-1}a)_2^+)$ for some $0 < j < n$ where $n = |\mathcal{O}_2(a)|$. Thus,

$$
\begin{aligned}
\mathrm{Hol}^\circ_{v_1}(\Phi)G_a(g) &= \varepsilon(h)Q_{a,\ldots,T_2^{j-1}a}(\varphi)G_a(g)\\
&= \varepsilon(h)Q_a(\varphi_{(1)})G_a(g)Q_{T_2 a}(\varphi_{(2)}) \cdots Q_{T_2^{j-1}a}(\varphi_{(j)})\\
&\overset{(3.7)}{=} \varepsilon(h)G_a(g_{(1)})Q_a(\varphi_{(1)} \leftharpoonup g_{(2)})Q_{T_2 a}(\varphi_{(2)}) \cdots Q_{T_2^{j-1}a}(\varphi_{(j)})\\
&= G_a(g_{(1)})\varepsilon(h)Q_{a,\ldots,T_2^{j-1}a}(\varphi \leftharpoonup g_{(2)})\\
&= G_a(g_{(1)})\,\mathrm{Hol}^\circ_{v_1}(\Phi \leftharpoonup (\varepsilon \otimes g_{(2)})),
\end{aligned}
$$

which, together with (7.13), is sufficient to conclude that (7.17) holds also if v_1 is a type 2 arc.

Case F_a with v_1 if v_1 is type 2. With the same v_1 as in the previous case, we can write

$$Q_{a,\ldots,T_2^{j-1}a}(\varphi)F_a(\psi) = Q_{a,\ldots,T_2^{j-1}a}(\varphi\psi_{(1)})Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(2)})$$
$$= Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(4)})Q_{a,\ldots,T_2^{j-1}a}(\psi_{(3)})Q_{a,\ldots,T_2^{j-1}a}(S(\psi_{(2)})\varphi\psi_{(1)})$$
$$= F_a(\psi_{(3)})Q_{a,\ldots,T_2^{j-1}a}(S(\psi_{(2)})\varphi\psi_{(1)}).$$

Case F_a with v_2 if v_1 is type 2. With v_1 as in the previous two cases the next arc is $v_2 = \big((T_2^j a)_0^-, (T_0^{-1}T_2^j a)_0^-,\ldots,(T_0^{-k+1}T_2^j a)_0^-\big)$, where $0 < k < |\mathcal{O}_0(T_2^j a)|$. Non-commutativity of F_a and $\mathrm{Hol}_{v_2}^\circ$ is due to the first arrow of v_2 and to the arrow $(T_2^j a)_2^+$ of γ_a,

$$P^S_{T_2^j a,\ldots,T_0^{-k+1}T_2^j a}(h)F_a(\psi) = P_{T_0^{-k+1}T_2^j a,\ldots,T_2^j a}(S(h))Q_{a,\ldots,T_2^{j-1}a}(\psi_{(1)})Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(2)})$$
$$= Q_{a,\ldots,T_2^{j-1}a}(\psi_{(1)})P_{T_0^{-k+1}T_2^j a,\ldots,T_0^{-1}T_2^j a}(S(h_{(2)}))P_{T_2^j a}(S(h_{(1)}))Q_{T_2^j a}(\psi_{(2)})$$
$$\times Q_{T_2^{j+1}a,\ldots,T_2^{n-1}a}(\psi_{(3)})$$
$$\overset{(3.4)}{=} Q_{a,\ldots,T_2^{j-1}a}(\psi_{(1)})P_{T_0^{-k+1}T_2^j a,\ldots,T_0^{-1}T_2^j a}(S(h_{(2)}))$$
$$\times Q_{T_2^j a}(\psi_{(2)} \leftharpoonup S(S(h_{(1)})_{(2)}))P_{T_2^j a}(S(h_{(1)})_{(1)})Q_{T_2^{j+1}a,\ldots,T_2^{n-1}a}(\psi_{(3)})$$
$$= Q_{a,\ldots,T_2^{j-1}a}(\psi_{(1)})Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(2)} \leftharpoonup h_{(1)})P_{T_0^{-k+1}T_2^j a,\ldots,T_2^j a}(S(h_{(2)})).$$

Combining this with the result of the previous case, we have

$$\mathrm{Hol}_{v_1 v_2}^\circ(\Phi)F_a(\psi) = \mathrm{Hol}_{v_1}^\circ\big(h_{(1)} \otimes \xi_i\varphi_{(1)}\xi_j\big)\mathrm{Hol}_{v_2}^\circ\big(S(x_j)h_{(2)}x_i \otimes \varphi_{(2)}\big)F_a(\psi)$$
$$= Q_{a,\ldots,T_2^{j-1}a}(\xi_i\varphi\xi_j)P^S_{T_2^j a,\ldots,T_0^{-k+1}T_2^j a}(S(x_j)hx_i)F_a(\psi)$$
$$= Q_{a,\ldots,T_2^{j-1}a}(\xi_i\varphi\xi_j)Q_{a,\ldots,T_2^{j-1}a}(\psi_{(1)})Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(2)} \leftharpoonup [S(x_j)hx_i]_{(1)})$$
$$\times P_{T_0^{-k+1}T_2^j a,\ldots,T_2^j a}\big(S([S(x_j)hx_i]_{(2)})\big)$$
$$= Q_{a,\ldots,T_2^{j-1}a}(\xi_i\varphi\xi_j\psi_{(1)})Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(3)})P^S_{T_2^j a,\ldots,T_0^{-k+1}T_2^j a}\big([S(x_j)hx_i] \leftharpoonup \psi_{(2)}\big)$$

The Hopf algebraic identity

$$\xi_i\varphi\xi_j\psi_{(1)} \otimes [S(x_j)hx_i] \leftharpoonup \psi_{(2)} = \psi_{(2)}\xi_i\varphi\xi_j \otimes S(x_j)(h \leftharpoonup \psi_{(1)})x_i$$

shows that the last line above equals to

$$Q_{a,\ldots,T_2^{j-1}a}(\psi_{(2)}\xi_i\varphi\xi_j)Q_{T_2^j a,\ldots,T_2^{n-1}a}(\psi_{(3)})P^S_{T_2^j a,\ldots,T_0^{-k+1}T_2^j a}(S(x_j)(h \leftharpoonup \psi_{(1)})x_i)$$
$$= F_a(\psi_{(2)})Q_{a,\ldots,T_2^{j-1}a}(\xi_i\varphi\xi_j)P^S_{T_2^j a,\ldots,T_0^{-k+1}T_2^j a}(S(x_j)(h \leftharpoonup \psi_{(1)})x_i)$$
$$= F_a(\psi_{(2)})\mathrm{Hol}_{v_1}^\circ\big((h \leftharpoonup \psi_{(1)})_{(1)} \otimes \xi_i\varphi_{(1)}\xi_j\big)\mathrm{Hol}_{v_2}^\circ\big(S(x_j)(h \leftharpoonup \psi_{(1)})_{(2)}x_i \otimes \varphi_{(2)}\big)$$
$$= F_a(\psi_{(2)})\mathrm{Hol}_{v_1 v_2}^\circ\big(h \leftharpoonup \psi_{(1)} \otimes \varphi\big)$$
$$= F_a(\psi_{(2)})\mathrm{Hol}_{v_1 v_2}^\circ\big(\Phi \leftharpoonup (\psi_{(1)} \otimes 1)\big).$$

Taking into account (7.16), this proves (7.18) for ρ having first arc of type 2.

The above cases together prove formula (7.11) for all proper left ribbons with disjoint source and target.

Formula (7.12) requires to consider another 6 cases but there is a shorter and more instructive way to prove it. We can apply duality to view the holonomy of the left ribbon ρ as the holonomy of a right ribbon in another Kitaev model based on the Hopf algebra H^*. This duality has two ingredients: an algebra isomorphism $I\colon \mathcal{M}(\Sigma, H) \to \mathcal{M}(\Sigma^*, H^*)$ and an isomorphism $J\colon \mathrm{Qv}(\mathcal{D}(P)^\sim) \to \mathrm{Qv}(\mathcal{D}(P^\smallfrown)^\sim)$ of quivers. The I is defined by $I(P_a(h)) := Q^*_{T_1 a}(h)$,

$I(Q_a(\varphi)) := P_a^*(\varphi)$ where the P_a^* and Q_a^* are the standard generators of $\mathcal{M}(\Sigma^*, H^*)$ like the P_a and Q_a have been for $\mathcal{M}(\Sigma, H)$ in Section 3. Comparing equations (3.7) and (3.18), it is easy to see that I is an algebra isomorphism. It maps Gauss' law operators to flux operators and vice versa,

$$I(G_a(h)) = F^*_{T_2^{-1}a}(h), \qquad I(F_a(\varphi)) = G^*_{T_2^{-1}a}(\varphi), \qquad I(\mathbf{D}_a(X)) = \mathbf{D}^*_{T_2^{-1}a}(\Upsilon(X)),$$

where

$$\Upsilon\colon \ \mathcal{D}(H) \to \mathcal{D}(H^*),$$
$$\varphi \otimes h \mapsto h_{(2)} \otimes \varphi_{(2)} \cdot \langle h_{(3)}, \varphi_{(1)} \rangle \langle S(h_{(1)}), \varphi_{(3)} \rangle$$

is an isomorphism of algebras and antiisomorphism of coalgebras.

In order to construct J, we first define a bijection j between the OCPMs associated to the arrow presentations P and its dual $P^\frown$ of (4.2) by

$$j\colon \ \Sigma(\mathrm{P}) \to \Sigma(\breve{\mathrm{P}}), \qquad \mathcal{O}_0(a) \mapsto \breve{\mathcal{O}}_2\,(T_1 a), \qquad \mathcal{O}_1(a) \mapsto \breve{\mathcal{O}}_1\,(a), \qquad \mathcal{O}_2(a) \mapsto \breve{\mathcal{O}}_0\,(a),$$

which satisfies

$$j \circ \partial_0 = \breve{d}_1, \qquad j \circ d_0 = \breve{\partial}_0, \qquad j \circ \partial_1 = \breve{d}_0, \qquad j \circ d_1 = \breve{\partial}_1\,.$$

The definition of J is this

$$J\colon \ \mathrm{Qv}(\mathcal{D}(\mathrm{P})^\sim) \to \mathrm{Qv}\big(\mathcal{D}(\breve{\mathrm{P}})^\sim\big), \qquad \langle v, f \rangle \mapsto \langle j(f), j(v) \rangle, \qquad a_0^\pm \mapsto (T_1 a)_2^\pm, \qquad a_2^\pm \mapsto a_0^\pm.$$

This can be shown to be a map of quivers by reading off the definition of ∇_i from (5.1). For example,

$$\breve{\nabla}_0 \,\circ J\big(a_0^+\big) = \breve{\nabla}_0\,(T_1 a)_2^+ = \big\langle \breve{\partial}_0\, T_1 a, \breve{d}_0\, T_1 a \big\rangle = \big\langle \breve{\partial}_1\, a, \breve{d}_1\, a \big\rangle$$
$$= \langle j(d_1 a), j(\partial_0 a) \rangle = J(\langle \partial_0 a, d_1 a \rangle) = J \circ \nabla_0\big(a_0^+\big).$$

This J is one of the maps that realizes the intuitively clear idea that the complexes $\mathcal{D}(\Sigma)^*$ and $\mathcal{D}(\Sigma^*)^*$ are the "same" except relabelling the arrows and repainting the blue and red faces to red and blue, respectively.

Now J induces an isomorphism $J^*\colon \mathsf{Path}\mathcal{D}(\mathrm{P})^\sim \to \mathsf{Path}\mathcal{D}(\breve{\mathrm{P}})^\sim$ of groupoids which sends a left ribbon $\rho\colon s(a) \to s(b)$ to a right ribbon $J^*\rho\colon \ \breve{s}\,(T_2^{-1}a) \to \breve{s}\,(T_2^{-1}b)$. (The appearance of T_2^{-1} is due to the fact that the site labelling function $a \mapsto s(a)$ satisfies $J \circ s = \breve{s} \circ T_2^{-1}$. As such, it belongs to the object map of the functor J^*.) If $\mathrm{Hol}^{\circ\frown}$ denotes the opholonomy of the dual model, then the duality formula says that for all curves ρ of the original model

$$I\big(\mathrm{Hol}_\rho^\circ(\Phi)\big) = \mathrm{Hol}_{J^*\rho}^{\circ\smile}(\Upsilon^*(\Phi)), \tag{7.19}$$

where $\Upsilon^* = \big(\Upsilon^\mathsf{T}\big)^{-1}\colon \mathcal{D}(H)^* \to \mathcal{D}(H^*)^*$ is the mapping

$$\Upsilon^*(h \otimes \varphi) = \xi_k \varphi \xi_l \otimes S(x_l)h x_k,$$

which is an antiisomorphism of algebras and an isomorphism of coalgebras.

Applying the duality formula to (7.11), we obtain

$$\mathrm{Hol}_{J^*\rho}^{\circ\smile}(\Upsilon^*(\Phi))\mathbf{D}^*_{T_2^{-1}a}(\Upsilon(X)) = \mathbf{D}^*_{T_2^{-1}a}\big(\Upsilon\big(X_{(1)}\big)\big)\,\mathrm{Hol}_{J^*\rho}^{\circ\smile}\big(\Upsilon^*(\Phi_{(2)})\big)\langle \Upsilon^*(\Phi_{(1)}), \Upsilon(X_{(2)}) \rangle$$
$$= \mathbf{D}^*_{T_2^{-1}a}\big(\Upsilon(X)_{(2)}\big)\,\mathrm{Hol}_{J^*\rho}^{\circ\smile}\big(\Upsilon^*(\Phi)_{(2)}\big)\langle \Upsilon^*(\Phi)_{(1)}, \Upsilon(X)_{(1)} \rangle$$
$$= \mathbf{D}^*_{T_2^{-1}a}\big(\Upsilon(X)_{(2)}\big)\,\mathrm{Hol}_{J^*\rho}^{\circ\smile}\big(\Upsilon^*(\Phi) \leftharpoonup \Upsilon(X)_{(1)}\big).$$

Therefore,

$$\mathrm{Hol}_\gamma^{\circ\frown}(\Psi)\mathbf{D}_a^*(Y) = \mathbf{D}_a^*\big(Y_{(2)}\big)\,\mathrm{Hol}_\gamma^{\circ\frown}\big(\Psi \leftharpoonup Y_{(1)}\big)$$

for all proper right ribbons $\gamma\colon \overset{\smile}{s}(a)\to\overset{\smile}{s}(b)$, for $Y\in\mathcal{D}(H^*)$ and for $\Psi\in\mathcal{D}(H^*)^*$. For the left ribbon γ^{-1}, this implies

$$\begin{aligned}
\mathbf{D}_a^*(Y)\,\mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}(\Psi) &= \mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}\big(\Psi_{(1)}\big)\,\mathrm{Hol}_\gamma^{\circ\frown}\big(\Psi_{(2)}\big)\mathbf{D}_a^*(Y)\,\mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}\big(\Psi_{(3)}\big) \\
&= \mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}\big(\Psi_{(1)}\big)\mathbf{D}_a^*\big(Y_{(2)}\big)\,\mathrm{Hol}_\gamma^{\circ\frown}\big(\Psi_{(2)}\leftharpoonup Y_{(1)}\big)\,\mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}\big(\Psi_{(3)}\big) \\
&= \mathrm{Hol}_{\gamma^{-1}}^{\circ\frown}\big(Y_{(1)}\rightharpoonup\Psi\big)\mathbf{D}_a^*\big(Y_{(2)}\big).
\end{aligned}$$

Since this is true for all left ribbons γ^{-1} with target $\overset{\smile}{s}(a)$ not only on $\mathcal{D}(\Sigma^*)^*$ and not only for H^* but on the dual of the double of any OCPM Σ and for any involutive f.d. Hopf algebra H, we can apply it to the original Σ and H and we get the proof of (7.12). ∎

For sake of completeness, we note here the formulas that hold instead of (7.9) and (7.10) if $\rho\colon s(a)\to s(b)$ is a proper right ribbon,

$$\mathrm{Hol}_\rho^\circ(\Phi)\underset{a}{\blacktriangleleft} X = \mathrm{Hol}_\rho^\circ(\Phi\leftharpoonup X), \tag{7.20}$$

$$X\underset{b}{\blacktriangleright}\mathrm{Hol}_\rho^\circ(\Phi) = \mathrm{Hol}_\rho^\circ(X\rightharpoonup\Phi). \tag{7.21}$$

Corollary 7.7. *Let $s(a)$ and $s(b)$ be disjoint sites and $\rho\colon s(a)\to s(b)$ be a proper left ribbon. Then the space $\mathrm{Hol}_\rho^\circ(\mathcal{D}(H)^*)$ of ribbon operators over ρ equipped with the actions $\underset{b}{\triangleright}$ and $\underset{a}{\triangleleft}$ is a bimodule over $\mathcal{D}(H)$ such that*

$$\mathrm{Hol}_\rho^\circ\colon\ \mathcal{D}(H)^*\to\mathrm{Hol}_\rho^\circ(\mathcal{D}(H)^*)$$

is a homomorphism of $\mathcal{D}(H)$-$\mathcal{D}(H)$ bimodules if we consider $\mathcal{D}(H)^*$ as the regular bimodule $\langle\mathcal{D}(H)^*,\rightharpoonup,\leftharpoonup\rangle$. This map is also a homomorphism of left $\mathcal{D}(H)$-module algebras and of right $\mathcal{D}(H)$-module algebras.

Similar holds for proper right ribbons $\rho\colon s(a)\to s(b)$ but then the bimodule structure is given by $\underset{b}{\blacktriangleright}$ and $\underset{a}{\blacktriangleleft}$ and the algebra structure by $\mathcal{D}(H)^{*\,\mathrm{op}}$.

Proof. That $\mathrm{Hol}_\rho^\circ(\mathcal{D}(H)^*)$ is a left and right $\mathcal{D}(H)$-module and that the map is a left and right $\mathcal{D}(H)$-module map follow from Theorem 7.6. Due to disjointness of the source and target sites, the left and right actions commute so $\mathrm{Hol}_\rho^\circ(\mathcal{D}(H)^*)$ is a bimodule. The last sentence follows from (7.4) which shows that $\mathrm{Hol}_\rho^\circ(\mathcal{D}(H)^*)\subset\mathcal{M}$ is a left and right module subalgebra. ∎

Similar results can be obtained [5] in a context of non-semisimple Hopf algebras.

We close the section with two more results on closed ribbon operators. An opcurve $\gamma\colon s_0\to s_1$ is called closed if $s_0=s_1$. If $\gamma=(d_1,\ldots,d_n)$ is closed, then every cyclic permutation $C^k\gamma=(d_{k+1},\ldots,d_n,d_1,\ldots,d_k)$ is also closed. It follows from the code word representation of ribbons that if the closed curve γ happens to be a ribbon then the $C^k\gamma$ are also ribbons of the same type (left/right). Similarly, the cyclic permutations of proper closed ribbons are again proper as it is clear from Definition 6.1.

Proposition 7.8. *Let $\delta\colon s(a)\to s(a)$ and $\rho\colon s(a)\to s(a)$ be closed left ribbons of the following form $\big(\text{for }\underset{U}{\leftrightarrow}\text{ see Definition 7.1}\big)$:*

$$\delta = \big(a_0^-\big)\delta'\big((T_2^{-1}a)_2^+\big), \qquad where \quad \delta'\underset{U}{\leftrightarrow}\rho,$$

$$\rho = \big(a_2^+\big)\rho'\big((T_0a)_0^-\big), \qquad where \quad \rho'\underset{U}{\leftrightarrow}\lambda,$$

so that the only intersection of δ and ρ is the site $s(a)$. (This is possible if the genus of $[\Sigma]$ is at least 1.) Then

$$\mathrm{Hol}^\circ_\delta(\Psi)\,\mathrm{Hol}^\circ_\rho(\Phi) = \mathrm{Hol}^\circ_\rho(R_{2'} \rightharpoonup \Phi \leftharpoonup R_1)\,\mathrm{Hol}^\circ_\delta(R_{1'} \rightharpoonup \Psi \leftharpoonup R_2) \tag{7.22}$$

for all $\Phi, \Psi \in \mathcal{D}(H)^*$ where $R = R_1 \otimes R_2 \equiv R_{1'} \otimes R_{2'}$ is the R-matrix of $\mathcal{D}(H)$.

Proof. For any curve γ we abbreviate $\mathrm{Hol}^\circ_\gamma$ by H_γ. The strategy of the proof is similar to that of Lemma 3.1. We decompose the opholonomies into 3 terms,

$$H_\delta(\Psi) = H_{(a_0^-)}\big(\Psi_{(1)}\big)\,H_{\delta'}\big(\Psi_{(2)}\big)\,H_{((T_2^{-1}a)_2^+)}\big(\Psi_{(3)}\big),$$
$$H_\rho(\Phi) = H_{(a_2^+)}\big(\Phi_{(1)}\big)\,H_{\rho'}\big(\Phi_{(2)}\big)\,H_{((T_0a)_0^-)}\big(\Phi_{(3)}\big)$$

with the middle terms commuting with all terms of the other line. Then we exchange H_δ and H_ρ in such a way that the order of terms of the same line remains intact in every step. Using the identities (6.5) and (6.6), we can write

$$\begin{aligned}
H_\delta(\Psi)H_\rho(\Phi) &= H_{(a_0^-)}\big(\Psi_{(1)}\big)H_{(a_2^+)}\big(\Phi_{(1)}\big)H_{\rho'}\big(\Phi_{(2)}\big) \\
&\quad \cdot H_{\delta'}\big(\Psi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(\Psi_{(3)}\big)H_{((T_0a)_0^-)}\big(\Phi_{(3)}\big) \\
&= H_{(a_2^+)}\big(\Phi_{(1)} \leftharpoonup R_1\big)H_{(a_0^-)}\big(\Psi_{(1)} \leftharpoonup R_2\big)H_{\rho'}\big(\Phi_{(2)}\big)H_{\delta'}\big(\Psi_{(2)}\big) \\
&\quad \cdot H_{((T_0a)_0^-)}\big(R_{2'} \rightharpoonup \Phi_{(3)}\big)H_{((T_2^{-1}a)_2^+)}\big(R_{1'} \rightharpoonup \Psi_{(3)}\big) \\
&= H_{(a_2^+)}\big(\Phi_{(1)} \leftharpoonup R_1\big)H_{\rho'}\big(\Phi_{(2)}\big)H_{((T_0a)_0^-)}\big(R_{2'} \rightharpoonup \Phi_{(3)}\big) \\
&\quad \cdot H_{(a_0^-)}\big(\Psi_{(1)} \leftharpoonup R_2\big)H_{\delta'}\big(\Psi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(R_{1'} \rightharpoonup \Psi_{(3)}\big) \\
&= H_\rho\big(R_{2'} \rightharpoonup \Phi \leftharpoonup R_1\big)H_\delta\big(R_{1'} \rightharpoonup \Psi \leftharpoonup R_2\big),
\end{aligned}$$

which completes the proof of (7.22). ∎

A closed ribbon version of (7.9) and (7.10) can be obtained as follows. First, we need a lemma.

Lemma 7.9. *Let $\rho\colon s(a) \to s(a)$ be a non-empty closed proper left ribbon which is not a face loop. Then there exist proper left ribbons $\rho_1\colon s(a) \to s(b)$ and $\rho_2\colon s(b) \to s(a)$ with disjoint $s(a)$ and $s(b)$ such that $\rho = \rho_1\rho_2$.*

Proof. Let $\rho = v_1 \cdots v_m$ be the arc decomposition as in the prof of Theorem 7.6. Since ρ is not a face loop, $m \geq 2$. Let $s(a_0) \xrightarrow{v_1} s(a_1) \xrightarrow{v_2} s(a_2)$, where $a_0 = a$. Then $s(a_1)$ is neighbour to $s(a_0)$ and $s(a_2)$ is disjoint from $s(a_0)$. Similarly, the sites of $s(a_{m-2}) \xrightarrow{v_{m-1}} s(a_{m-1}) \xrightarrow{v_m} s(a_m)$, where $a_m = a$, satisfy that $s(a_{m-1})$ is neighbour to $s(a_m)$ and $s(a_{m-2})$ is disjoint from $s(a_m)$. This already excludes $m = 2$ and $m = 3$. So we can take $\rho_1 = v_1v_2$ and $\rho_2 = v_3 \cdots v_m$. ∎

Proposition 7.10. *Let $\rho\colon s(a) \to s(a)$ be a non-empty closed proper left ribbon. Then*

$$\mathbf{D}_a(X)\,\mathrm{Hol}_\rho(\Phi) = \mathrm{Hol}_\rho\big(X_{(1)} \rightharpoonup \Phi \leftharpoonup S_\mathcal{D}(X_{(3)})\big)\mathbf{D}_a\big(X_{(2)}\big) \tag{7.23}$$

holds true for every $X \in \mathcal{D}(H)$ and $\Phi \in \mathcal{D}(H)^*$.

Proof. First, we assume that ρ is not a face loop. Take a decomposition

$$\rho = s(a) \xrightarrow{\rho_1} s(b) \xrightarrow{\rho_2} s(a)$$

provided by Lemma 7.9 with disjoint sites $s(a)$ and $s(b)$. Using Theorem 7.6,

$$
\begin{aligned}
\mathbf{D}_a(X)\operatorname{Hol}_\rho(\Phi) &= \mathbf{D}_a(X)\operatorname{Hol}_{\rho_1}\big(\Phi_{(1)}\big)\operatorname{Hol}_{\rho_2}\big(\Phi_{(2)}\big)\\
&= \operatorname{Hol}_{\rho_1}\big(\Phi_{(1)}\leftharpoonup S_{\mathcal D}(X_{(2)})\big)\mathbf{D}_a\big(X_{(1)}\big)\operatorname{Hol}_\rho\big(\Phi_{(2)}\big)\\
&= \operatorname{Hol}_{\rho_1}\big(\Phi_{(1)}\leftharpoonup S_{\mathcal D}(X_{(3)})\big)\operatorname{Hol}_\rho\big(X_{(1)}\rightharpoonup\Phi_{(2)}\big)\mathbf{D}_a\big(X_{(2)}\big)\\
&= \operatorname{Hol}_\rho\big(X_{(1)}\rightharpoonup\Phi\leftharpoonup S_{\mathcal D}(X_{(3)})\big)\mathbf{D}_a\big(X_{(2)}\big)
\end{aligned}
$$

the statement is proven.

If ρ is a face loop, then ρ is either γ_a or α_a^{-1}. Assume $\rho=\gamma_a$. Then by Proposition 6.5

$$
\operatorname{Hol}_\rho(h\otimes\varphi)=\varepsilon(h)F_a(\varphi)=\mathbf{D}_a(\varphi\otimes 1)\,\varepsilon(h),
$$

so the left-hand side of (7.23) with $X=\psi\otimes g$ and $\Phi=h\otimes\varphi$ can be written as

$$
\begin{aligned}
\mathbf{D}_a(X)\operatorname{Hol}_\rho(\Phi) &= D_a(X(\varphi\otimes 1)\varepsilon(h)\\
&= \mathbf{D}_a\big(\psi\varphi_{(2)}\otimes g_{(2)}\big)\cdot\big\langle\varphi_{(3)},g_{(1)}\big\rangle\big\langle\varphi_{(1)},S\big(g_{(3)}\big)\big\rangle\varepsilon(h).
\end{aligned}
$$

For the right-hand side, a much longer computation is necessary,

$$
\begin{aligned}
&\operatorname{Hol}_\rho\big(X_{(1)}\rightharpoonup\Phi\leftharpoonup S_{\mathcal D}(X_{(3)})\big)\mathbf{D}_a\big(X_{(2)}\big)\\
&= \operatorname{Hol}_\rho\big(\big[\underbrace{S(x_j)h_{(2)}x_i}_{h_{ji}}\otimes\varphi_{(2)}\big]_{(1)}\big)\big\langle\big[S(x_j)h_{(2)}x_i\otimes\varphi_{(2)}\big]_{(2)},X_{(1)}\big\rangle\\
&\quad\times\big\langle h_{(1)}\otimes\xi_i\varphi_{(1)}\xi_j,S_{\mathcal D}\big(X_{(3)}\big)\big\rangle\mathbf{D}_a\big(X_{(2)}\big)\\
&= \operatorname{Hol}_\rho\big(h_{ji(1)}\otimes\xi_k\varphi_{(2)}\xi_l\big)\big\langle S(x_l)h_{ji(2)}x_k\otimes\varphi_{(3)},X_{(1)}\big\rangle\\
&\quad\times\big\langle h_{(1)}\otimes\xi_i\varphi_{(1)}\xi_j,S_{\mathcal D}\big(X_{(3)}\big)\big\rangle\mathbf{D}_a\big(X_{(2)}\big)\\
&= \mathbf{D}_a\big(\xi_k\varphi_{(2)}\xi_l\otimes 1\big)\big\langle S(x_l)h_{ji}x_k\otimes\varphi_{(3)},\psi_{(3)}\otimes g_{(1)}\big\rangle\\
&\quad\times\big\langle x_nS\big(h_{(1)}\big)S(x_m)\otimes\xi_mS(\xi_j)S\big(\varphi_{(1)}\big)S(\xi_i)\xi_n,\psi_{(1)}\otimes g_{(3)}\big\rangle\mathbf{D}_a\big(\psi_{(2)}\otimes g_{(2)}\big)\\
&= \big\langle\psi_{(3)},S(x_l)S(x_j)h_{(2)}x_ix_k\big\rangle\big\langle\varphi_{(3)},g_{(1)}\big\rangle\big\langle\psi_{(1)},x_nS\big(h_{(1)}\big)S(x_m)\big\rangle\\
&\quad\times\big\langle\xi_mS(\xi_j)S\big(\varphi_{(1)}\big)S(\xi_i)\xi_n,g_{(3)}\big\rangle\mathbf{D}_a\big(\xi_k\varphi_{(2)}\xi_l\psi_{(2)}\otimes g_{(2)}\big)\\
&= \big\langle\psi_{(7)},h_{(2)}\big\rangle\big\langle\varphi_{(3)},g_{(1)}\big\rangle\big\langle\psi_{(2)},S\big(h_{(1)}\big)\big\rangle\\
&\quad\times\big\langle S\big(\psi_{(3)}\big)\psi_{(6)}S\big(\varphi_{(1)}\big)S\big(\psi_{(8)}\big)\psi_{(1)},g_{(3)}\big\rangle\mathbf{D}_a\big(\psi_{(9)}\varphi_{(2)}S\big(\psi_{(5)}\big)\psi_{(4)}\otimes g_{(2)}\big)\\
&= \varepsilon(h)\big\langle\varphi_{(3)},g_{(1)}\big\rangle\big\langle S\big(\varphi_{(1)}\big),g_{(3)}\big\rangle\mathbf{D}_a\big(\psi\varphi_{(2)}\otimes g_{(2)}\big).
\end{aligned}
$$

The case of $\rho=\alpha_a^{-1}$ can be proven by a similar calculation.   ∎

Note that (7.23) is identical to the multiplication rule in the double of the double, $\mathcal{D}(\mathcal{D}(H))$.

8   Homotopy

In order to explain "topological invariance" of the Kitaev model, the first step is to look for an appropriate notion of homotopy of curves. Eventually, we want to use homotopy on the complex $\mathcal{D}(\Sigma)^*$ but the very notion should be given for an arbitrary Σ.

Definition 8.1. Let Σ be a connected OCPM with arrow presentation $\langle\operatorname{Arr},T_0,T_2\rangle$.

(i) Let f be a face of Σ. An f-loop λ is a closed opcurve $(a,T_2a,T_2^2a,\dots)$ or its inverse where $a\in\operatorname{Arr}$ is such that $\mathcal{O}_2(a)=f$. The set of f-loops for all $f\in\Sigma^2$ are called face loops.

(ii) A lasso ϑ is a conjugate of a face loop, i.e., $\vartheta = \sigma\lambda\sigma^{-1}$, where λ is a face loop and σ is arbitrary, provided $\partial_1\sigma = \partial_0\lambda$. If λ is an f-loop then ϑ is called an f-lasso.

(iii) Let α, β be parallel opcurves, i.e., $\partial_i\alpha = \partial_i\beta$, $i = 0, 1$. A (combinatorial) homotopy from α to β is a sequence of lassos $(\vartheta_N, \ldots, \vartheta_1)$ such that

$$\beta\vartheta_N \ldots \vartheta_2\vartheta_1 = \alpha \qquad \text{as paths.}$$

Two opcurves are called homotopic, denoted by $\alpha \sim \beta$, if they are parallel and there is a homotopy from one of them to the other.

(iv) A contraction of a closed curve $\alpha\colon v \to v$ is a homotopy from α to the trivial curve (v). A closed curve is contractible if has a contraction.

(v) If $(\vartheta_N, \ldots, \vartheta_1)\colon \alpha \to \beta$ is a homotopy with ϑ_i being an f_i-lasso, then the support of this homotopy is the set of faces occurring in the list (f_i).

Remark 8.2. There are op-versions of the above definition of homotopy of opcurves:

- In ophomotopy of opcurves $\alpha \to \beta$ means $\alpha\vartheta_1 \cdots \vartheta_N = \beta$.

- In homotopy of curves $\alpha \to \beta$ means $\vartheta_N \circ \cdots \circ \vartheta_1 \circ \alpha = \beta$.

- In ophomotopy of curves $\alpha \to \beta$ means $\vartheta_1 \circ \cdots \circ \vartheta_N \circ \beta = \alpha$.

Scholium 8.3. Let Σ be a connected OCPM. Homotopy of opcurves on Σ satisfies the following elementary properties:

(i) If $\boldsymbol{\vartheta} = (\vartheta_N, \ldots, \vartheta_1)$ is a homotopy $\alpha \to \beta$ then $\boldsymbol{\vartheta}^{-1} = (\vartheta_1^{-1}, \ldots, \vartheta_N^{-1})$ is a homotopy $\beta \to \alpha$.

(ii) If $\alpha \xrightarrow{\boldsymbol{\vartheta}} \beta \xrightarrow{\boldsymbol{\vartheta'}} \gamma$ are homotopies, then the composition

$$\boldsymbol{\vartheta'} \circ \boldsymbol{\vartheta} := (\vartheta'_M, \ldots, \vartheta'_1, \vartheta_N, \ldots, \vartheta_1)$$

is a homotopy $\alpha \to \gamma$.

(iii) If two opcurves are equal as paths, then they are homotopic.

(iv) Homotopy is an equivalence relation on the set of opcurves and each path is the subset of a unique homotopy class.

(v) A homotopy $\alpha \to \beta$ is the same as a contraction of $\beta^{-1}\alpha$.

(vi) If $\boldsymbol{\vartheta}\colon \alpha \to \beta$ is a homotopy, then $(\beta\vartheta_N\beta^{-1}, \ldots \beta\vartheta_1\beta^{-1})$ is a homotopy $\beta^{-1} \to \alpha^{-1}$.

(vii) Let

$$u \xrightarrow{\alpha} v \underset{\gamma}{\overset{\beta}{\rightrightarrows}} w \xrightarrow{\delta} z$$

be opcurves and let $\boldsymbol{\vartheta}\colon \beta \to \gamma$ be a homotopy. Then $(\ldots, \delta^{-1}\vartheta_i\delta, \ldots)$ is a homotopy $\alpha\beta\delta \to \alpha\gamma\delta$.

(viii) If $(\vartheta_N, \ldots, \vartheta_1)\colon \alpha \to \beta$ is a homotopy and $p \in S_N$ is any permutation, then there exist closed curves π_i such that $\vartheta'_i := \pi_i\vartheta_{p(i)}\pi_i^{-1}$ defines a homotopy $(\vartheta'_N, \ldots, \vartheta'_1)\colon \alpha \to \beta$.

The set of homotopy classes of opcurves together with concatenation is a groupoid which is the opposite of the traditional fundamental groupoid $\pi_1(\Sigma)$. It is evident from Definition 8.1, although its formalization could be tedious, that this groupoid is the same as the fundamental groupoid $\pi_1([\Sigma])$ of the geometric realization of the complex Σ.

The notion of contraction defined above does not reflect the idea that a contractible curve bounds a disk and the support of the contraction is the disk itself. As a matter of fact, any
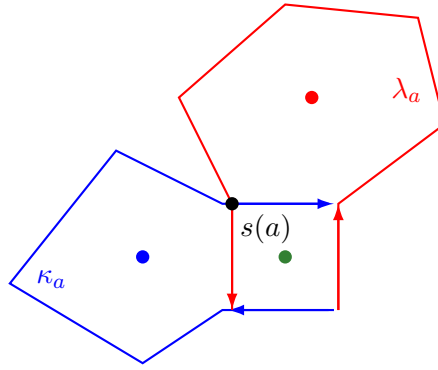
Figure 6. The basic curves of ribbon homotopy.

finite set of faces can be the support of a contraction. So one should study simple curves and their simple homotopies. But before we would dig too deeply into the subject of lasso homotopy we should remind ourselves of our main goal which is the algebra of ribbon operators on $\mathcal{D}(\Sigma)^*$. It turns out that the use of lassos in holonomy operators is extremely inconvenient. The lassos $\sigma\lambda\sigma^{-1}$ are not proper ribbons and the long tails σ introduce serious non-commutativities which practically obstructs any computation. We need some other kind of homotopy specialized to ribbon curves on $\mathcal{D}(\Sigma)^*$ by means of which we can deform (proper) ribbons in such a way that the curve stays a (proper) ribbon in each step.

Let us introduce the following open "lassos" (Figure 6):

$$\kappa_a := \left((T_1 a)_2^+, \left(T_0^{m-1} a \right)_0^-, \ldots, (T_0 a)_0^-, a_2^+ \right) : \; s(T_1 a) \to s(T_2 a), \tag{8.1}$$

$$\lambda_a := \left((T_1 a)_0^-, (T_2 a)_2^+, \ldots, \left(T_2^{n-1} a \right)_2^+, a_0^- \right) : \; s(T_1 a) \to s\left(T_0^{-1} a \right), \tag{8.2}$$

where $m = |\mathcal{O}_0(a)|$ and $n = |\mathcal{O}_2(a)|$. Note that $\kappa_a, \lambda_a \in \mathsf{Ribb}_L^\circ$.

Unlike the lassos the κ-s and λ-s cannot be nicely composed. So the strategy in ribbon curve homotopy is not composing the elementary paths, it is more like the procedure of grafting trees in which κ_a and λ_a play the role of the grafts.

Remark 8.4. Speaking about subcurves and subpaths, in general, requires some care. For opcurves σ and ρ, the relation $\sigma \subset \rho$ means the existence of opcurves ρ_1 and ρ_2 such that $\rho = \rho_1 \sigma \rho_2$ as opcurves, hence no reduction is allowed. For paths of opcurves, $\sigma \subset \rho$ means the existence of paths ρ_1 and ρ_2 such that $\rho = \rho_1 \sigma \rho_2$ as paths. In case of both σ and ρ are ribbons (of the same type), the relation $\sigma \subset \rho$ for the ribbon opcurves is stronger than the same relation for their paths. However, if we think more categorically and utilize the fact that $\mathsf{Ribb}_{L/R}^\circ$ is a category for concatenation, then we can use the notion of subribbon as a third possibility. For left ribbons, e.g., $\sigma \subset \rho$ is a subribbon if there exist left ribbons ρ_1 and ρ_2 such that $\rho = \rho_1 \sigma \rho_2$ either as opcurves or as paths because concatenation of left ribbons is always reduced so these two alternatives are equivalent. This third version of "subcurve" is which we need in the next definition.

Definition 8.5.

(a) For left ribbon opcurves $\rho, \rho' : s_0 \to s_1$, we say that the pair (ρ, ρ') is an elementary contraction and that (ρ', ρ) is an elementary relaxation in the following 4 cases:

- Vertex contraction: For some $a \in \mathrm{Arr}(\Sigma)$ the ρ contains a subribbon α_a^{-1} (see (6.7)) and ρ' is obtained from ρ by discarding this α_a^{-1}.
- Face contraction: For some $a \in \mathrm{Arr}(\Sigma)$, the ρ contains a subribbon γ_a (see (6.9)) and ρ' is obtained from ρ by discarding this γ_a.

- κ-contraction: For some $a \in \mathrm{Arr}(\Sigma)$, the ρ contains a subribbon $\kappa_a \subset \rho$ and ρ' is obtained by replacing this subribbon by $\big((T_1 a)_0^-\big)$.

- λ-contraction: For some $a \in \mathrm{Arr}(\Sigma)$, the ρ contains a subribbon $\lambda_a \subset \rho$ and ρ' is obtained by replacing this subribbon by $\big((T_1 a)_2^+\big)$.

For $\rho, \rho' \in \mathsf{Ribb}_L^\circ$, we write $\rho \approx \rho'$ and say that they are ribbon-homotopic if there is a sequence of left ribbons $(\rho_0, \ldots, \rho_k)$, the homotopy, such that $\rho_0 = \rho$, $\rho_k = \rho'$ and each move (ρ_i, ρ_{i+1}) is an elementary contraction or an elementary relaxation.

(b) For closed ribbons, we also define free ribbon-homotopy as a ribbon-homotopy in which we allow also circular moves: (ρ_i, ρ_{i+1}) is a circular move if ρ_{i+1} differs from ρ_i only in the choice of the base point. Free ribbon-homotopy is denoted by $\overset{\star}{\approx}$.

(c) For right ribbons, the definitions (a) and (b) can be used without change except that in the definition of moves the curves α_a^{-1}, γ_a, κ_a, λ_a have to be replaced by α_a, γ_a^{-1}, κ_a^{-1}, λ_a^{-1}, respectively.

(d) The support of a (free) ribbon-homotopy is the set $F \subseteq \Sigma$ of faces of $\mathcal{D}(\Sigma)^*$ arising as the union of the supports of the elementary contractions/relaxations. In case of vertex or face contraction/relaxation, the support is just the vertex or face in question. In case of κ-contraction/relaxation, the support has 2 elements, a vertex and an edge of Σ, the ones being winded around by κ_a. For λ-contraction/relaxation, the support is a similarly defined pair of a face and of an edge.

Notice that we have not defined homotopy between a left and a right ribbon.

In order to place ribbon-homotopy into the wider context of homotopy, we need to see abundance of ribbons in some sense.

Proposition 8.6. *Let Σ be a connected OCPM and let $s, s' \in \mathcal{D}(\Sigma)^{*0}$ be two sites.*

(i) *There exists a left ribbon $\rho \colon s \to s'$.*

(ii) *If $\gamma \colon s \to s'$ is an opcurve, then there exists a left ribbon $\rho \colon s \to s'$ such that $\gamma \sim \rho$, i.e., γ and ρ are homotopic in the sense of Definition 8.1.*

(iii) *If $\rho_1, \rho_2 \colon s \to s'$ are left ribbons such that $\rho_1 \approx \rho_2$, then $\rho_1 \sim \rho_2$.*

Proof. (ii) Let γ be the composite $s = s(a_0) \xrightarrow{c_1} s(a_1) \xrightarrow{c_2} \cdots \xrightarrow{c_n} s(a_n) = s'$, $v_k = \mathcal{O}_0(a_k)$, $f_k = \mathcal{O}_2(a_k)$ and let $w = W_n \cdots W_1$ be the code word of γ. The proof follows the principle of replacing bad letters with good words. Construct the new word

$$w' := w'_n \cdots w'_1, \qquad \text{where} \quad w'_k := \begin{cases} W_k & \text{if } W_k \in \{T_0^{-1}, T_2\}, \\ T_0^{-|v_k|+1} & \text{if } W_k = T_0, \\ T_2^{|f_k|-1} & \text{if } W_k = T_2^{-1}, \end{cases}$$

and define the ribbons $\rho_k := \langle a_0, w'_k \cdots w'_1 \rangle$. Then $\vartheta_k := \rho_{k-1}(c_k)\rho_k^{-1}$ is a lasso such that $\vartheta_1 \vartheta_2 \cdots \vartheta_n = \gamma \rho_n^{-1}$. Therefore, $\rho = \rho_n$ is the desired ribbon and $\big(\gamma^{-1}\vartheta_1 \gamma, \ldots, \gamma^{-1}\vartheta_n \gamma\big) \colon \gamma \to \rho$ is the desired homotopy.

(i) Since $\mathcal{D}(\Sigma)^*$ is connected, there exists an opcurve $\gamma = s \to s'$. Thus, (i) follows from (ii).

(iii) It suffices to show that the elementary relaxations are homotopies. For the cases of vertex and face relaxations, this is obvious. For κ_a, we can easily check the equality of paths $\big((T_1 a)_0^-\big) = \kappa_a \vartheta_v \vartheta_e$ with lassos $\vartheta_v = (a_2^-)\alpha_a(a_2^+)$ and $\vartheta_e = (a_2^-, a_0^-)\beta_a^{-1}(a_0^+, a_2^+)$. Thus, $(\vartheta_v, \vartheta_e) \colon \kappa_a \to \big((T_1 a)_0^-\big)$ is a homotopy. Similar formulas can be found for λ_a. ∎

Assume that we are given a closed ribbon which is contractible, e.g., because we know that the ribbon bounds a disk. How can we design a ribbon-homotopy which shrinks the ribbon to zero? This is the kind of problem in which "tree curves" play a role.

Definition 8.7. Tree curves or, more descriptively, tree dressing curves are closed ribbon opcurves occurring in two types:

A tree curve τ_F can be constructed from a tree $F \subset \Sigma^0 \cup \Sigma^1$, as a subgraph of Σ, as follows. Considering F as a set of faces of $\mathcal{D}(\Sigma)^*$ we take τ_F to be any of the negatively oriented reduced boundary opcurves of this fattened tree F.

A dual tree curve τ_F can be constructed from a dual tree $F \subset \Sigma^2 \cup \Sigma^1$, as a subgraph of Σ^*, as follows. Consider F as a set of faces of $\mathcal{D}(\Sigma)^*$ and take τ_F to be any of the positively oriented reduced boundary opcurves of this fattened tree F.

Predecessors of the tree curves are the "dual blocks" and "direct blocks" of Bombin and Martin-Delgado [1]. We need also rooted versions of them which are open ribbons:

Definition 8.8. A rooted tree curve is obtained from a tree curve τ_F by discarding the maximal arc running around a leaf vertex $v \in F$. This curve is denoted by $\tau_{F'}$, where $F' = F \setminus \{v\}$ is the underlying rooted tree. The root of $\tau_{F'}$ is the arrow which would complete the discarded arc to a full loop around v. If r is the root, (r) is called the root curve.

A rooted dual tree curve is obtained from a dual tree curve τ_F by discarding the maximal arc running around a leaf face $f \in F$. This curve is denoted by $\tau_{F'}$, where $F' = F \setminus \{f\}$ is the underlying rooted dual tree. The root of $\tau_{F'}$ is the arrow which would complete the discarded arc to a full loop around f. If r is the root, (r) is called the root curve.

Note the deviation from the usual notion of rooted tree. Here the root vertex is not an arbitrary vertex but a leaf. Furthermore, the root vertex does not belong to the rooted tree, so in fact there is only a root edge which is characterized by having only one boundary vertex (a "half edge").

The inverses τ_F^{-1} $\left(\tau_{F'}^{-1}\right)$ of the above (rooted) tree curves are needed in ribbon-homotopy of right ribbons, so they should also be called (rooted) tree curves.

Lemma 8.9. *The following facts are either obvious or easy to prove:*

(i) *For every (rooted) (dual) tree F, the τ_F is a proper left ribbon and τ_F^{-1} is a proper right ribbon.*

(ii) *If F is a singleton tree in Σ (in Σ^*), then the (dual) tree curve τ_F is α_a^{-1} (γ_a) for some $a \in \mathrm{Arr}(\Sigma)$.*

(iii) *If the rooted tree F' in Σ (in Σ^*) consists of 2 elements, then the rooted (dual) tree curve $\tau_{F'}$ is a κ_a (λ_a) for some $a \in \mathrm{Arr}(\Sigma)$.*

(iv) *Every (dual) tree curve can be obtained from an α_a^{-1} (γ_a) by successive grafting of κ_a-s (λ_a-s) by their roots to some arrows of color T_0^{-1} (T_2). The same can be said about the rooted (dual) tree curves if we start with a κ (λ).*

(v) *Every κ_a, λ_a is ribbon-homotopic to its root curve.*

(vi) *Every tree curve and dual tree curve is ribbon-homotopic to some trivial ribbon.*

(vii) *Every rooted tree curve and rooted dual tree curve is ribbon-homotopic to its root curve.*

Under (iv) the "grafting" means a κ- or λ-relaxation under the condition that the result remains a proper opcurve. We do not have to use unproper relaxations to get tree curves although a general ribbon-homotopy would allow them.
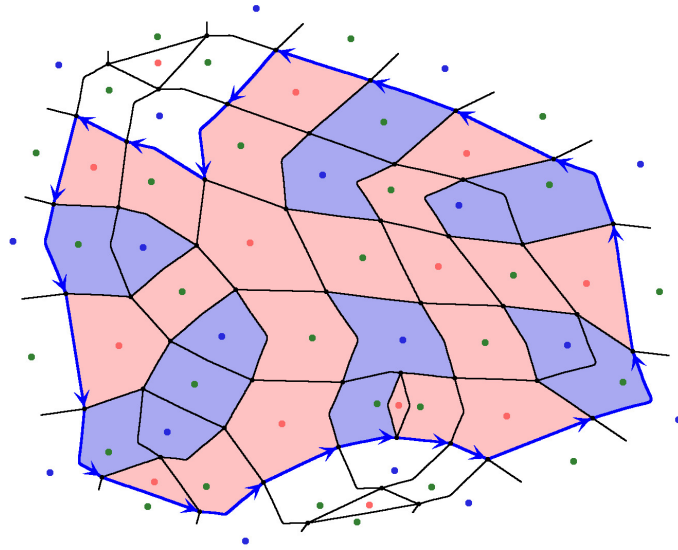
Figure 7. A plan for contracting a left ribbon. (See the proof of Theorem 8.10.)

Theorem 8.10. *For a connected OCPM Σ, let $\rho\colon s \to s$ be a closed proper left ribbon. We divide the faces of $\mathcal{D}(\Sigma)^*$ into connected components by saying that two faces p and q of $\mathcal{D}(\Sigma)^*$ are connected if there is a dual curve $\delta\colon p \to q$ on $\mathcal{D}(\Sigma)^*$ such that neither the arrows of δ nor their opposites are members of ρ. Assume that*

- *ρ cuts the set of faces into 2 connected components, L consisting of the faces lying on the left and R consisting of the faces lying on the right of ρ;*

- *as subcomplexes of Σ either L or R has Euler characteristic 1.*

Then there is a ribbon-homotopy from ρ to the trivial ribbon (s) such that the support of the homotopy is L, resp. R, and in each step of the homotopy the curve is proper.

Proof. Looking at the two connected components L and R as sets of cells of Σ, we have $L \cap R = \varnothing$ and $L \cup R = \Sigma$. But there is a substantial difference between these subcomplexes. L is upper semiclosed, i.e., together with any cell c it contains also $\mathrm{Cb}(c)$, while R is lower semiclosed[2], i.e., together with any cell c it contains also $\mathrm{Bd}(c)$. This is visible on Figure 7 by the appearance of alternating red and green dots on the left-hand side of ρ and of the blue and green dots on the right-hand side. These sequences of faces and edges on the left and vertices and edges on the right are the boundary dual curve of L and the boundary curve of R, respectively. They coincide with the two borderlines $E^*(\rho) = \langle c, b \rangle$ of ρ, see in and below Lemma 5.4. Since ρ is a ribbon, these borderlines never cross by Lemma 5.8.

Let $L^d = \Sigma^d \cap L$ for $d = 0, 1, 2$ and assume that it is L which has Euler characteristic $\chi(L) = 1$. Choose a maximal dual tree $F \subset L^2 \cup L^1$ in the connected complex L. Let τ_F be its dual tree curve the base point of which is that of ρ. Since F contains all faces of L but none of its vertices and F has $\chi(F) = 1$, the complement $G := L \setminus F$ has Euler characteristic

$$\chi(G) = \left| G^0 \right| - \left| G^1 \right| = \left| L^0 \right| - \left(\left| L^1 \right| - \left(\left| L^2 \right| - 1 \right) \right) = \left| L^0 \right| - \left| L^1 \right| + \left| L^2 \right| - 1 = \chi(L) - 1 = 0.$$

As a subgraph of the planar L the G cannot contain circles because they would separate some faces from the maximal dual tree F. Therefore, G is a disjoint union of trees $G = G_1 \cup \cdots \cup G_k$ each of which has to be rooted in order to satisfy $\chi(G_i) = 0$. The root edge of G_i, in the

---

[2]A subset of a connected CPM Σ which is both lower and upper semiclosed is either empty or the whole Σ.
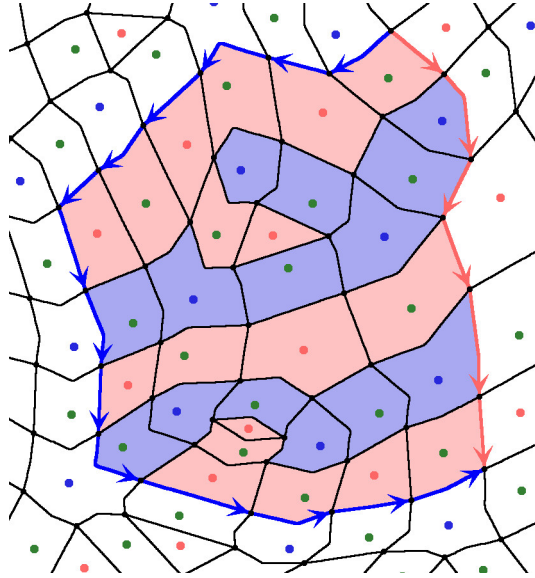
Figure 8. The plan of a proper ribbon-homotopy from the blue to the red left ribbon. (See the proof of Theorem 8.13.)

language of $\mathcal{D}(\Sigma)^*$, is a type 1 face e_i having one arrow $r_i \in \rho$ of color T_0^{-1} on its boundary. This is the root of G_i. Starting with the root r_i, we can perform a sequence of κ-relaxations along the tree G_i. Doing this for all G_i, we arrive to that ρ is a ribbon-homotopic to τ_F. Since τ_F can be contracted by λ-contractions to the trivial curve, we are done. During each steps of the homotopy we had only proper ribbons and used only grafts κ_a, λ_b which belonged entirely to L. Furthermore, the supports of the used grafts form a partition of L into vertex-edge pairs and face-edge pairs, so the support of the homotopy is precisely L. (In this case, we call L the contraction disk of ρ.)

If it is R which has $\chi(R) = 1$, then the proof goes by constructing a maximal tree $F \subset R^0 \cup R^1$ and the forest of rooted dual trees $G = R \setminus F$. Then the homotopy first λ-relaxes ρ inside R along the forest and then κ-contracts the resulting tree curve τ_F. ∎

The special ribbon-homotopies we found in the above theorem turn out to be the only ribbon-homotopies applicable in the Kitaev model. So we give them a name.

Definition 8.11. A (free) ribbon-homotopy $(\rho_0, \rho_1, \ldots, \rho_n)$ is called a proper (free) ribbon-homotopy if each ρ_i is proper (Definition 6.1). A proper ribbon which has a proper ribbon-homotopy to a trivial curve is called properly contractible.

Example 8.12. The figure 8 opcurve $8_a := \alpha_a^{-1} \gamma_a$ is a left ribbon which is not proper. It has a ribbon-homotopy to $(s(a))$ but no proper ribbon-homotopy at all. The same can be said about the double loop $\gamma_a \gamma_a$.

Theorem 8.13. *Let $\rho_1, \rho_2 \colon s(a) \to s(b)$ be parallel non-trivial proper left ribbons such that the closed opcurve $\rho_1 \rho_2^{-1}$ is the positively oriented boundary curve of a disk $B \subset \mathcal{D}(\Sigma)^{*2}$. Then there is a proper ribbon-homotopy $\rho_1 \to \rho_2$ of support B.*

Proof. By assumption, B lies on the left of ρ_1 and the right of ρ_2. The B is now neither upper nor lower semiclosed and, strangely enough, has Euler characteristic 0 in Σ. This is an artifact of the different behaviour at the two borderline segments ρ_1 and ρ_2: B looks like upper semiclosed along the ρ_1 and lower semiclosed along ρ_2. So we can start growing rooted trees (with roots on ρ_1 necessarily) until we get a maximal forest $F \subset B$ of rooted trees (see
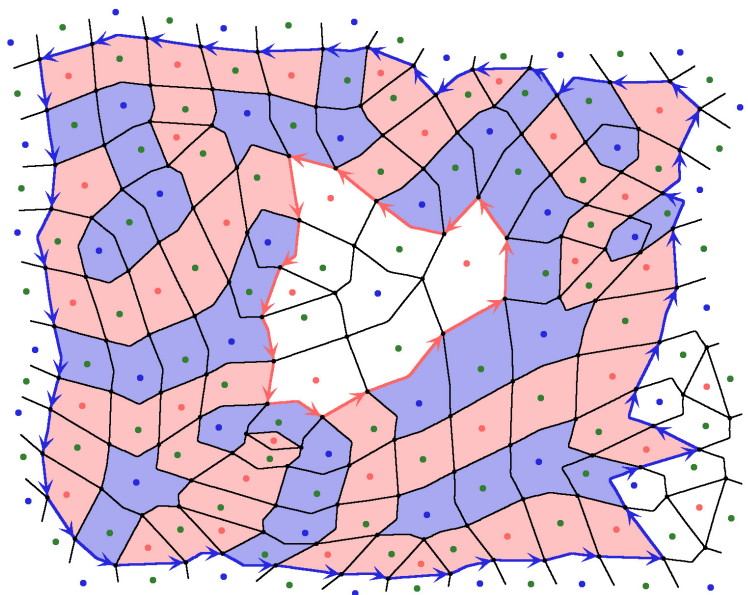
Figure 9. For a free proper ribbon-homotopy from the blue ribbon to the red ribbon first choose a maximal forest of rooted trees on the annulus. These are the bluish faces. Then perform κ-relaxations on the blue ribbon along the bluish faces. What we get is a Hamilton cycle; a closed proper ribbon (not shown) visiting every site of the annulus exactly once. Then we make a circular move along the Hamilton cycle in order to transplant the base point from the blue ribbon to the red ribbon. Finally, we perform λ-contractions along the reddish faces.

Figure 8). The complement $F^* = B \setminus F$ is then a maximal forest of rooted dual trees (with roots on ρ_2 necessarily). The homotopy then starts with κ-relaxations of ρ_1 along the trees of F. The result is a proper left ribbon $\gamma\colon s(a) \to s(b)$ which consists of the (broken) borderline between F and F^* completed by those arrows of ρ_1 and ρ_2 which are not roots of some tree. Finally, a sequence of λ-contractions applied to γ along F^* results in the ribbon ρ_2.   ■

The intermediate curve γ in the proof of the above theorem has the remarkable property of passing through every site of the disk B exactly once. So it is a Hamiltonian opcurve on the graph $\downarrow B \setminus B$ obtained by taking the lower semiclosed closure of B in the complex $\mathcal{D}(\Sigma)^*$ and then discarding all its faces. The existence of such Hamiltonian curve is crucial in the following free ribbon-homotopy.

Theorem 8.14. *Let A be a finite connected set of faces of $\mathcal{D}(\Sigma)^*$ with Euler characteristic 0 in Σ such that the positively oriented boundary of A in $\mathcal{D}(\Sigma)^*$ consists of two components, a closed proper left ribbon ρ_1 and a closed proper right ribbon ρ_2^{-1}. Then there is a proper free ribbon-homotopy $\rho_1 \to \rho_2$ supported on the annulus A.*

Proof. The proof is very similar to the proof of Theorem 8.13 therefore we content ourselves with referring to Figure 9.   ■

The annulus of Figure 9 is obtained by taking the relative complement of a disk within a larger disk. But this need not be so. The proof of Theorem 8.14 nowhere uses that the "hole" is a disk. So it applies to any annulus strip winding around a torus or on some handle of a higher genus surface.

9   Ground states

So far the Kitaev model was treated on a pure algebraic level. Now we turn it into a model in quantum field theory. So we assume that the ground field $K = \mathbb{C}$, the field of complex numbers, and the finite-dimensional Hopf algebra H is a C*-Hopf algebra. This means, on the one hand, that H is endowed with an antilinear involution, called the star operation,[3] satisfying

$$(gh)^\dagger = h^\dagger g^\dagger, \quad h_{(1)}{}^\dagger \otimes h_{(2)}{}^\dagger = \left(h^\dagger\right)_{(1)} \otimes \left(h^\dagger\right)_{(2)}, \qquad g, h \in H.$$

Consequently,

$$1^\dagger = 1, \qquad \varepsilon\!\left(h^\dagger\right) = \overline{\varepsilon(h)}, \qquad S(h)^\dagger = S^{-1}\!\left(h^\dagger\right), \qquad i^\dagger = i.$$

On the other hand, one requires that H has a faithful star-representation on a Hilbert space, i.e., such that $h^\dagger$ is represented by the adjoint of the operator representing h. From now on H denotes a finite-dimensional C*-Hopf algebra. Finite-dimensional C*-Hopf algebras are semisimple and have $S^2 = \mathrm{id}$. So every Hopf algebraic relations used before remain valid. In particular, H^* has a Haar integral ι and the formula

$$(g, h) := \left\langle \iota, g^\dagger h \right\rangle$$

is a scalar product making H a Hilbert space and the left regular representation a faithful star representation, $(k, hg) = \left(h^\dagger k, g\right)$.

The dual Hopf algebra H^* is then also a C*-Hopf algebra. The star operation of H^* is given by

$$\left\langle \varphi^\dagger, h \right\rangle := \overline{\left\langle \varphi, S(h)^\dagger \right\rangle}, \qquad \varphi \in H^*, \quad h \in H.$$

It follows that the double $\mathcal{D}(H)$ and its dual $\mathcal{D}(H)^*$ are both C*-Hopf algebras with respective star operations

$$(\psi \otimes g)^\dagger = \psi^\dagger_{(2)} \otimes g^\dagger_{(2)} \cdot \overline{\left\langle \psi_{(3)}, S\!\left(g_{(1)}\right) \right\rangle} \, \overline{\left\langle \psi_{(1)}, g_{(3)} \right\rangle}, \qquad (h \otimes \varphi)^\dagger = h^\dagger \otimes \varphi^\dagger.$$

Assume Σ is a finite connected OCPM. Since $\mathcal{M}(\Sigma)$ is a full matrix algebra, it has, up to isomorphisms, only one irreducible representation. The standard choice for the Hilbert space of this representation is $\mathscr{H}(\Sigma) := H^{\otimes \Sigma^1}$ which is the tensor product of copies H_e of the Hilbert space H for each edge $e \in \Sigma^1$. The action of $\mathcal{M}$ on $\mathscr{H}$ is defined by means of an orientation of the edges, i.e., by a section $e \mapsto a_e$ of the function $\mathcal{O}_1 \colon \mathrm{Arr}(\Sigma) \to \Sigma^1$. Namely, $P_a(h)$ and $Q_a(\varphi)$ act as the identity on each tensor factor H_e, $e \neq \mathcal{O}_1(a)$, and for $e = \mathcal{O}_1(a)$ the $P_a(h)$ is left multiplication by h and $Q_a(\varphi)$ is $\varphi \rightharpoonup \_$ on H_e if $a = a_e$ and $P_a(h)$ is right multiplication by $S(h)$ and $Q_a(\varphi) = \_ \leftharpoonup S(\varphi)$ on H_e if $a = T_1 a_e$; cf. (3.10) and (3.11). In this way the abstract operators $P_a(h)$, $Q_a(\varphi)$ become concrete operators on the Hilbert space $\mathscr{H}$ such that $P_a(h)^\dagger = P_a\!\left(h^\dagger\right)$ and $Q_a(\varphi)^\dagger = Q_a\!\left(\varphi^\dagger\right)$, where $(\cdot)^\dagger$ on the right-hand side is the abstract star operation while on the left-hand side it is the adjoint of an operator on $\mathscr{H}$.

The adjoints of one-step holonomies (6.3) satisfy $\mathrm{Hol}^\circ_{(d)}(\Psi)^\dagger = \mathrm{Hol}^\circ_{(d)}\!\left(\Psi^\dagger\right)$. Therefore, if γ is an arbitrary opcurve, then

$$\mathrm{Hol}^\circ_\gamma(\Psi)^\dagger = \mathrm{Hol}^\circ_{\gamma^{-1}}\!\left(S_{\mathcal{D}^*}(\Psi)^\dagger\right), \qquad \Phi \in \mathcal{D}(H)^*. \tag{9.1}$$

But if the curve is a proper ribbon, then we also have

$$\mathrm{Hol}^\circ_\rho(\Psi)^\dagger = \mathrm{Hol}^\circ_\rho\!\left(\Psi^\dagger\right), \qquad \Psi \in \mathcal{D}(H)^*, \quad \text{if } \rho \text{ is proper.} \tag{9.2}$$

---

[3]Having already many stars in this paper we shall denote the star operation, as well as the adjoints of bounded linear operators on a Hilbert space, by $(\cdot)^\dagger$.

In particular, the Gauss' law and flux operators satisfy $G_a(h)^\dagger = G_a(h^\dagger)$ and $F_a(\varphi)^\dagger = F_a(\varphi^\dagger)$. So A_v and B_f are self adjoint idempotents, i.e., projectors, and the Hamiltonian (3.32) is self adjoint, too.

The ground states of the Hamiltonian are represented by vectors Ω in $\mathscr{H}$ such that

$$A_v\Omega = \Omega \qquad \text{and} \qquad B_f\Omega = \Omega \qquad \forall v \in \Sigma^0, \quad f \in \Sigma^2.$$

Following [12], we denote by $\mathscr{P}$ the subspace in $\mathscr{H}$ of such vacuum vectors Ω and call it the protected space. Explicit wave functions of vacuum vectors have been calculated in [3] using a tensor network formalism. If H is a group algebra of a finite group G, then a basis of $\mathscr{P}$ consists of gauge equivalence classes of G-valued flat gauge fields on Σ [5]. So the dimension of $\mathscr{P}$ depends only on the genus of Σ through the fundamental group $\pi_1(\Sigma)$. A far reaching generalization of the "protected object" $\mathscr{P}$ appeared recently in [10] in which the Kitaev quantum double model is studied over Hopf monoids in monoidal categories.

In terms of expectation values $\omega(M) = (\Omega, M\Omega)$, the ground states can be abstractly characterized by the equations

$$\omega(MA_v) = \omega(M) = \omega(MB_f), \qquad M \in \mathcal{M}, \quad v \in \Sigma^0, \quad f \in \Sigma^2$$

for the normalized ($\omega(\mathbf{1}) = 1$) positive linear functional $\omega\colon \mathcal{M} \to \mathbb{C}$. Knowing ω, which we do not, is complete information about the representation on $\mathscr{H}$ by the Gelfand–Neumark–Segal construction. So the representation of $\mathcal{M}$ on $\mathscr{H}$ may be called the vacuum representation.

The above properties of the vacua $\Omega \in \mathscr{P}$ imply

$$\mathrm{Hol}^\circ_{\alpha_a}(\Phi)\Omega = \varphi(1)G_a(h)G_a(i)\Omega = \varphi(1)\varepsilon(h)\Omega = \Omega \cdot \varepsilon_{\mathcal{D}^*}(\Phi), \tag{9.3}$$

$$\mathrm{Hol}^\circ_{\gamma_a}(\Phi)\Omega = \varepsilon(h)F_a(\varphi)F_a(\iota)\Omega = \varepsilon(h)\varphi(1)\Omega = \Omega \cdot \varepsilon_{\mathcal{D}^*}(\Phi), \tag{9.4}$$

hence also the formula expressing gauge invariance of the vacua

$$\mathbf{D}_a(X)\Omega = \Omega \cdot \varepsilon_D(X), \qquad a \in \mathrm{Arr}(\Sigma), \quad X \in \mathcal{D}(H), \quad \Omega \in \mathscr{P}.$$

Proposition 9.1. *Let $a \in \mathrm{Arr}(\Sigma)$, $v = \mathcal{O}_0(a)$, $f = \mathcal{O}_2(a)$ and let $\Omega \in \mathscr{H}$ be any vector. Then for all $\Phi \in \mathcal{D}(H)^*$*

$$\mathrm{Hol}^\circ_{\alpha_a^{\pm 1}}(\Phi)\Omega = \Omega \cdot \varepsilon_{\mathcal{D}^*}(\Phi) \qquad \text{whenever} \quad A_v\Omega = \Omega, \tag{9.5}$$

$$\mathrm{Hol}^\circ_{\gamma_a^{\pm 1}}(\Phi)\Omega = \Omega \cdot \varepsilon_{\mathcal{D}^*}(\Phi) \qquad \text{whenever} \quad B_f\Omega = \Omega, \tag{9.6}$$

$$\mathrm{Hol}^\circ_{\kappa_a}(\Phi)\Omega = \mathrm{Hol}^\circ_{((T_1a)_0^-)}\Omega \qquad \text{whenever} \quad A_v\Omega = \Omega, \tag{9.7}$$

$$\mathrm{Hol}^\circ_{\lambda_a}(\Phi)\Omega = \mathrm{Hol}^\circ_{((T_1a)_2^+)}\Omega \qquad \text{whenever} \quad B_f\Omega = \Omega, \tag{9.8}$$

$$\mathrm{Hol}^\circ_{\kappa_a^{-1}}(\Phi)\Omega = \mathrm{Hol}^\circ_{((T_1a)_0^+)}\Omega \qquad \text{whenever} \quad A_v\Omega = \Omega, \tag{9.9}$$

$$\mathrm{Hol}^\circ_{\lambda_a^{-1}}(\Phi)\Omega = \mathrm{Hol}^\circ_{((T_1a)_2^-)}\Omega \qquad \text{whenever} \quad B_f\Omega = \Omega, \tag{9.10}$$

where κ_a, λ_a are the left ribbons defined in (8.1) and (8.2).

Proof. Formulas (9.5) and (9.6) can be proven as in (9.3) and (9.4). The proof of (9.7) goes as follows. κ_a contains α_a^{-1} except one arrow, $(T_0^m a)_0^- = a_0^-$. So we write its opholonomy as

$$\mathrm{Hol}^\circ_{\kappa_a}(\Phi) = \mathrm{Hol}^\circ_{((T_1a)_2^+, a_0^+)}(\Phi_{(1)})\, \mathrm{Hol}^\circ_{\alpha_a^{-1}}(\Phi_{(2)})\, \mathrm{Hol}^\circ_{(a_2^+)}(\Phi_{(3)}).$$

We leave the first term for a while and concentrate on the action of the rest on the vacuum:

$$\mathrm{Hol}^\circ_{\alpha_a^{-1}}(\Phi_{(1)})\, \mathrm{Hol}^\circ_{(a_2^+)}(\Phi_{(2)})\Omega = G_a(S(h_{(1)}))\langle\xi_i\varphi_{(1)}\xi_j, 1\rangle\varepsilon(S(x_j)h_{(2)}x_i)Q_a(\varphi_{(2)})\Omega$$

$$= G_a(S(h))Q_a(\varphi)\Omega \overset{(3.23)}{=} Q_a(\varphi \leftharpoonup h_{(1)})G_a(S(h_{(2)}))\Omega$$

$$= Q_a(\varphi \leftharpoonup h)\Omega,$$

where, as usual, $\Phi = h \otimes \varphi$ stands for an arbitrary element of $\mathcal{D}(H)^*$. In order to express this result in terms of Φ, we remark that

$$u = \xi_i \otimes S(x_i) \qquad \text{and} \qquad u^{-1} = \xi_i \otimes x_i.$$

Therefore,

$$
\begin{aligned}
u^{-1} \rightharpoonup \Phi &= \big(h_{(1)} \otimes \xi_j \varphi_{(1)} \xi_k\big) \cdot \big\langle S(x_k) h_{(2)} x_j \otimes \varphi_{(2)}, \xi_i \otimes x_i \big\rangle \\
&= \big(h_{(1)} \otimes \xi_j \varphi_{(1)} \xi_k\big) \cdot \big\langle \varphi_{(2)}, S(x_k) h_{(2)} x_j \big\rangle \\
&= h_{(1)} \otimes \varphi_{(4)} \varphi_{(1)} S\big(\varphi_{(2)}\big) \cdot \big\langle \varphi_{(3)}, h_{(2)} \big\rangle \\
&= h_{(1)} \otimes \varphi \leftharpoonup h_{(2)}
\end{aligned}
$$

and $\varphi \leftharpoonup h = (\varepsilon \otimes \mathrm{id})\big(u^{-1} \rightharpoonup \Phi\big)$, so we have

$$Q_a(\varphi \leftharpoonup h) = \mathrm{Hol}^{\circ}_{(a_2^+)}\big(u^{-1} \rightharpoonup \Phi\big).$$

In this way, we have proven that

$$
\begin{aligned}
\mathrm{Hol}^{\circ}_{\kappa_a}(\Phi)\Omega &= \mathrm{Hol}^{\circ}_{((T_1 a)_2^+, a_0^+)}\big(\Phi_{(1)}\big) \, \mathrm{Hol}^{\circ}_{(a_2^+)}\big(u^{-1} \rightharpoonup \Phi_{(2)}\big)\Omega = \\
&= \mathrm{Hol}^{\circ}_{((T_1 a)_2^+, a_0^+, a_2^+, (T_1 a)_0^+)}\big(\Phi_{(1)}\big) \, \mathrm{Hol}^{\circ}_{((T_1 a)_0^-)}\big(u^{-1} \rightharpoonup \Phi_{(2)}\big)\Omega.
\end{aligned}
$$

Now the opcurve in the first term is a conjugate of the central curve β_a which, by (6.10) and Lemma 6.7, we know to have opholonomy equal to $\mathbf{1}_{\mathcal{M}} \cdot \langle \_, S_{\mathcal{D}}(u)\rangle$. Thus,

$$\mathrm{Hol}^{\circ}_{\kappa_a}(\Phi)\Omega = \mathrm{Hol}^{\circ}_{((T_1 a)_0^-)}\big(S_{\mathcal{D}}(u) u^{-1} \rightharpoonup \Phi\big)\Omega.$$

For C^*-Hopf algebras, $S_{\mathcal{D}}(u) = u$, therefore (9.7) is proven.

The proof of (9.8) goes as follows. We insert $\big(a_2^-, a_2^+\big)$ into λ_a in order to complete the incomplete face loop,

$$\mathrm{Hol}^{\circ}_{\lambda_a}(\Phi) = \mathrm{Hol}^{\circ}_{((T_1 a)_0^-, a_2^-)}\big(\Phi_{(1)}\big) \, \mathrm{Hol}^{\circ}_{\gamma_a}\big(\Phi_{(2)}\big) \, \mathrm{Hol}^{\circ}_{(a_0^-)}\big(\Phi_{(3)}\big).$$

Next we compute the action of the last two terms on the vacuum by using the (inverse of the) exchange relation (3.29),

$$
\begin{aligned}
\mathrm{Hol}^{\circ}_{\gamma_a}\big(\Phi_{(1)}\big) \, \mathrm{Hol}^{\circ}_{(a_0^-)}\big(\Phi_{(2)}\big)\Omega &= \varepsilon\big(h_{(1)}\big) F_a\big(\xi_i \varphi_{(1)} \xi_j\big) P_a\big(S\big(S(x_j) h_{(2)} x_i\big)\big)\varphi_{(2)}(1) \\
&= P_a\big((\xi_i \varphi \xi_j)_{(1)} \rightharpoonup S(x_i) S(h) x_j\big) F_a\big((\xi_i \varphi \xi_j)_{(2)}\big)\Omega \\
&= P_a(\xi_i \varphi \xi_j \rightharpoonup S(x_i) S(h) x_j)\Omega.
\end{aligned}
$$

Next we simplify the last expression by proving the identity

$$
\begin{aligned}
\xi_i \varphi \xi_j \rightharpoonup S(x_i) S(h) x_j &= \big\langle \Phi_{(1)}, \varepsilon \otimes x_i \big\rangle \big\langle \Phi_{(2)}, \xi_j \otimes 1 \big\rangle \cdot (\xi_i \rightharpoonup S(x_j)) \\
&= \big\langle \Phi, (\varepsilon \otimes x_i)(\xi_j \xi_k \otimes 1) \big\rangle S(x_k) \langle \xi_i, S(x_j)\rangle \\
&= \big\langle \Phi, \underbrace{(\varepsilon \otimes S(x_j)(\xi_j \otimes 1)}_{S_{\mathcal{D}}(R_1) R_2 = S_{\mathcal{D}}(u)}(\xi_k \otimes 1)\big\rangle S(x_k) \\
&= S((\mathrm{id} \otimes \varepsilon)(\Phi \leftharpoonup S_{\mathcal{D}}(u))),
\end{aligned}
$$

which implies

$$P_a(\xi_i \varphi \xi_j \rightharpoonup S(x_i) S(h) x_j) = \mathrm{Hol}^{\circ}_{(a_0^-)}(\Phi \leftharpoonup S_{\mathcal{D}}(u)).$$

Putting together and using that $S_\mathcal{D}(u)$ is central, we obtain

$$\begin{aligned}
\mathrm{Hol}^\circ_{\lambda_a}(\Phi)\Omega &= \mathrm{Hol}^\circ_{(T_1 a)^-_0, a^-_2}\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_{(a^-_0)}\big(\Phi_{(2)} \leftharpoonup S_\mathcal{D}(u)\big)\Omega \\
&= \mathrm{Hol}^\circ_{(T_1 a)^-_0, a^-_2, a^-_0}\big(\Phi_{(1)}\big)\big\langle \Phi_{(2)}, S_\mathcal{D}(u)\big\rangle \\
&= \mathrm{Hol}^\circ_{(T_1 a)^-_0, a^-_2, a^-_0, (T_1 a)^-_2}\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_{((T_1 a)^+_2)}\big(S_\mathcal{D}(u) \rightharpoonup \Phi_{(2)}\big).
\end{aligned}$$

Now the opcurve in the first term is a conjugate of β_a^{-1} therefore it is central and its opholonomy is $\mathbf{1}_\mathcal{M} \cdot \big\langle \,_-, S_\mathcal{D}(u)^{-1}\big\rangle$ by Lemma 6.7. This immediately implies (9.8).

Since κ_a, λ_a are proper ribbons, (9.9) and (9.10) for all Φ are equivalent to (9.7) and (9.8), respectively, by (7.5). ∎

As a generalization of the protected space, let us introduce the spaces $\mathscr{P}(F) \subseteq \mathscr{H}$ of partial vacua as follows. Consider $F \subseteq \Sigma$ as a set of faces of $\mathcal{D}(\Sigma)^*$ and let $F^d := F \cap \Sigma^d$. Then define

$$\mathscr{P}(F) := \big\{\Omega \in \mathscr{H} \mid A_v\Omega = \Omega,\, B_f\Omega = \Omega,\text{ for } v \in F^0,\, f \in F^2\big\}.$$

So in particular $\mathscr{P}(\Sigma) = \mathscr{P}$ and $\mathscr{P}(\varnothing) = \mathscr{H}$.

Theorem 9.2 (topological invariance). *For a connected OCPM Σ, let ρ_1, ρ_2 be proper ribbons on $\mathcal{D}(\Sigma)^*$ of the same left/right type and let $F \subset \mathcal{D}(\Sigma)^{*2}$. Then*

(i) *if there is a proper ribbon-homotopy $\rho_1 \to \rho_2$ of support F, then*

$$\mathrm{Hol}^\circ_{\rho_1}(\Phi)\Omega = \mathrm{Hol}^\circ_{\rho_2}(\Phi)\Omega \qquad \text{for all} \quad \Phi \in \mathcal{D}(H)^*, \quad \Omega \in \mathscr{P}(F);$$

(ii) *if ρ_1, ρ_2 are closed and there is a proper free ribbon-homotopy $\rho_1 \to \rho_2$ of support F then*

$$\mathrm{Hol}^\circ_{\rho_1}(\Psi)\Omega = \mathrm{Hol}^\circ_{\rho_2}(\Psi)\Omega \qquad \text{for all} \quad \Psi \in \mathrm{Cocom}\,\mathcal{D}(H)^*, \quad \Omega \in \mathscr{P}(F).$$

Proof. (i) It suffices to consider elementary ribbon-homotopies (ρ_1, ρ_2) in which a subribbon $\sigma_1 \subset \gamma\sigma_1\delta = \rho_1$ is replaced by the subribbon $\sigma_2 \subset \gamma\sigma_2\delta = \rho_2$. (So $\sigma_1 \mapsto \sigma_2$ is either a vertex or face or κ- or λ-contraction/relaxation.) Then the argument is this

$$\begin{aligned}
\mathrm{Hol}^\circ_{\rho_1}(\Phi)\Omega &= \mathrm{Hol}^\circ_\gamma\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_{\sigma_1}\big(\Phi_{(2)}\big)\,\mathrm{Hol}^\circ_\delta\big(\Phi_{(3)}\big)\Omega = [\rho_1\text{ is proper}] \\
&= \mathrm{Hol}^\circ_\gamma\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_\delta\big(\Phi_{(3)}\big)\,\mathrm{Hol}^\circ_{\sigma_1}\big(\Phi_{(2)}\big)\Omega = [\text{Proposition } 9.1] \\
&= \mathrm{Hol}^\circ_\gamma\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_\delta\big(\Phi_{(3)}\big)\,\mathrm{Hol}^\circ_{\sigma_2}\big(\Phi_{(2)}\big)\Omega = [\rho_2\text{ is proper}] \\
&= \mathrm{Hol}^\circ_\gamma\big(\Phi_{(1)}\big)\,\mathrm{Hol}^\circ_{\sigma_2}\big(\Phi_{(2)}\big)\,\mathrm{Hol}^\circ_\delta\big(\Phi_{(3)}\big)\Omega = \mathrm{Hol}^\circ_{\rho_2}(\Phi)\Omega.
\end{aligned}$$

(ii) follows from (i) if we add the observation that for cocommutative Ψ we can cyclically permute proper ribbons already at the level of operators: $\mathrm{Hol}^\circ_{\rho_1}(\Psi) = \mathrm{Hol}^\circ_{\rho_2}(\Psi)$ whenever (ρ_1, ρ_2) is a circular move, see (7.6). ∎

Corollary 9.3. *Let γ be a closed proper ribbon, $\rho\colon \langle v_0, f_0\rangle \to \langle v_1, f_1\rangle$ any proper ribbon and let $\Psi \in \mathrm{Cocom}\,\mathcal{D}(H)^*$ and $\Phi \in \mathcal{D}(H)^*$.*

(i) *If $\Omega \in \mathscr{P}$, then $\mathrm{Hol}^\circ_\gamma(\Psi)\Omega$ also belongs to $\mathscr{P}$ and depends only on the proper free ribbon-homotopy class of γ.*

(ii) *If γ is properly contractible (Theorem 8.10) and $\Omega \in \mathscr{P}$, then $\mathrm{Hol}^\circ_\gamma(\Psi)\Omega = \Omega \cdot \varepsilon_{\mathcal{D}^*}(\Psi)$.*

(iii) *If $\Omega \in \mathscr{P}$, then $\mathrm{Hol}^\circ_\rho(\Phi)\Omega$ belongs to $\mathscr{P}(\Sigma \setminus \{v_0, v_1, f_0, f_1\})$ and depends only on the proper ribbon-homotopy class of ρ.*

Proof. This is a direct consequence of Corollary 7.5 and Theorem 9.2 with the additional remark that in (iii) the support of any proper ribbon-homotopy $\rho \to \rho'$ is automatically disjoint from $\{v_0, v_1, f_0, f_1\}$. ∎

The natural question that emerges is whether the proper ribbon-homotopy classes are the same as the ribbon-homotopy classes or the homotopy classes. If Σ is finite and has genus $\mathbf{g} > 0$, then the answer is no. While the latter two are infinite groupoids there are only finitely many proper ribbons on a finite Σ. In addition, proper ribbons cannot be composed so the proper ribbon-homotopy classes need not even form a groupoid. On the infinite plane, however, one expects that every ribbon-homotopy class contains a proper ribbon so all the three should reduce to the pair groupoid on the set of sites.

10   Excited states

As explained in [12], the holonomy of an open ribbon $\rho\colon s_0 \to s_1$ can be thought to create some charge at s_1 and the anticharge at s_0. These charges can then be measured by contractible closed ribbon operators winding around s_i. We can now make this precise in the Hopf algebraic Kitaev model.

Theorem 10.1. *Let Σ be a connected OCPM. On $\mathcal{D}(\Sigma)^*$ let $\rho\colon s_0 \to s_1$ be a proper left ribbon and $\gamma\colon s \to s$ a closed proper left ribbon.*

(*i*) *We assume that γ is properly contractible in the sense of Theorem 8.10 so divides $\mathcal{D}(\Sigma)^{*2}$ into two connected components L and R among which now L is the disk.*

(*ii*) *L contains both faces of the site $s_1 = \langle v_1, f_1 \rangle$.*

(*iii*) *R contains both faces of the site $s_0 = \langle v_0, f_0 \rangle$.*

Then for all partial vacuum vectors $\Omega \in \mathscr{P}(L)$

$$\mathrm{Hol}_\gamma^\circ(\Psi)\,\mathrm{Hol}_\rho^\circ(\Phi)\Omega = \mathrm{Hol}_\rho^\circ(\overset{\leftrightarrow}{\Psi} \rightharpoonup \Phi)\Omega \tag{10.1}$$

for all $\Phi \in \mathcal{D}(H)^*$ and $\Psi \in \mathrm{Cocom}\,\mathcal{D}(H)^*$ where the linear isomorphism $\mathcal{D}(H)^* \to \mathcal{D}(H)$, $\Psi = k \otimes \psi \mapsto \overset{\leftrightarrow}{\Psi} := \psi \otimes k$ maps $\mathrm{Cocom}\,\mathcal{D}(H)^*$ onto $\mathrm{Center}\,\mathcal{D}(H)$.

Proof. L is upper semiclosed and contains v_1 and f_1. The smallest upper semiclosed subcomplex of Σ which contains v_1 and f_1 is nothing but $\uparrow v_1$ consisting[4] of v_1 and of its vertex neighbourhood $\mathrm{Nb}(v_1)$ (Definition 2.3). So $\uparrow v_1 \subseteq L$ and its boundary left ribbon δ is proper. The set A of cells between δ and γ is an annulus as in Theorem 8.14 so there is a proper free ribbon-homotopy $\gamma \to \delta$ of support A. Since $L \supseteq A$, $\Omega \in \mathscr{P}(L) \subseteq \mathscr{P}(A)$ and by Corollary 7.5 also $\mathrm{Hol}_\rho^\circ(\Phi)\Omega \in \mathscr{P}(A)$. Using topological invariance in the sense of Theorem 9.2 (ii), we have

$$\mathrm{Hol}_\gamma^\circ(\Psi)\,\mathrm{Hol}_\rho^\circ(\Phi)\Omega = \mathrm{Hol}_\delta^\circ(\Psi)\,\mathrm{Hol}_\rho^\circ(\Phi)\Omega.$$

Now using topological invariance, this time in the sense of Theorem 9.2 (i), of $\mathrm{Hol}_\rho^\circ(\Phi)\Omega$, we can choose ρ to be as simple inside δ as possible. By proper ribbon-homotopy supported in L, we

---

[4]An example of a $\uparrow v$ is the hole in the annulus of Figure 9.

can deform ρ in such a way that only the last arrow of the new ρ lies
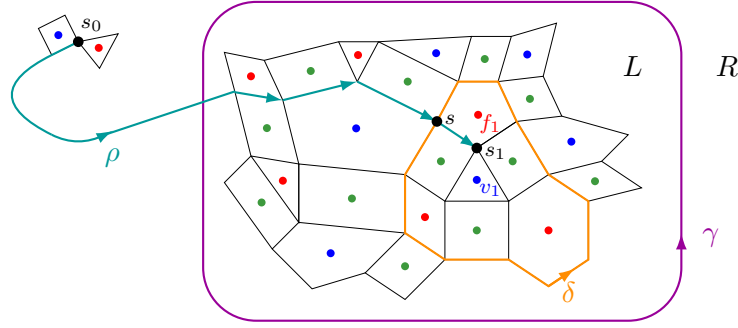
inside δ and the last two arrows of this ρ form a straight curve intersecting δ only at a single site. So we can assume that ρ and δ have the following form:

$$\rho = \rho'\big((T_0 a)_0^-, a_2^+\big), \qquad \delta = \big(a_0^-\big)\delta'\big((T_2^{-1}a)_2^+\big),$$

where a is chosen such that $s(a)$ is the intersection point of ρ and δ, hence $s_1 = s(T_2 a)$. Notice that we have chosen the base point of δ to be the intersection point. The ρ' and δ' are proper ribbons with holonomies that commute with that of all other pieces of opcurves. Now we can compute the exchange relation of $H_\delta(\Psi) = \mathrm{Hol}_\delta^\circ(\Psi)$ and $H_\rho(\Phi) = \mathrm{Hol}_\rho^\circ(\Phi)$ as follows:

$$\begin{aligned}
H_\delta(\Psi)H_\rho(\Phi) &= H_{(a_0^-)}\big(\Psi_{(1)}\big)H_{\delta'}\big(\Psi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(\Psi_{(3)}\big) \\
&\quad \times H_{\rho'}\big(\Phi_{(1)}\big)H_{((T_0 a)_0^-)}\big(\Phi_{(2)}\big)H_{(a_2^+)}\big(\Phi_{(3)}\big) \\
&\overset{(6.6)}{=} H_{\rho'}\big(\Phi_{(1)}\big)H_{(a_0^-)}\big(\Psi_{(1)}\big)H_{\delta'}\big(\Psi_{(2)}\big) \\
&\quad \times H_{((T_0 a)_0^-)}\big(R_2 \rightharpoonup \Phi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(R_1 \rightharpoonup \Psi_{(3)}\big)H_{(a_2^+)}\big(\Phi_{(3)}\big) \\
&= H_{\rho'}\big(\Phi_{(1)}\big)H_{((T_0 a)_0^-)}\big(R_2 \rightharpoonup \Phi_{(2)}\big)H_{(a_0^-)}\big(\Psi_{(1)}\big)H_{(a_2^+)}\big(\Phi_{(3)}\big) \\
&\quad \times H_{\delta'}\big(\Psi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(R_1 \rightharpoonup \Psi_{(3)}\big) \\
&\overset{(6.5)}{=} H_{\rho'}\big(\Phi_{(1)}\big)H_{((T_0 a)_0^-)}\big(R_2 \rightharpoonup \Phi_{(2)}\big)H_{(a_2^+)}\big(\Phi_{(3)} \leftharpoonup R_{1'}\big)H_{(a_0^-)}\big(\Psi_{(1)} \leftharpoonup R_{2'}\big) \\
&\quad \times H_{\delta'}\big(\Psi_{(2)}\big)H_{((T_2^{-1}a)_2^+)}\big(R_1 \rightharpoonup \Psi_{(3)}\big) \\
&= H_{\rho'}\big(\Phi_{(1)}\big)H_{((T_0 a)_0^-)}\big(\Phi_{(2)}\big)\cdot\big\langle\Phi_{(3)}, R_2 R_{1'}\big\rangle\cdot H_{(a_2^+)}\big(\Phi_{(4)}\big)\cdot H_\delta(R_1 \rightharpoonup \Psi \leftharpoonup R_{2'}).
\end{aligned}$$

The result is reminiscent of (7.22) but now, since the endpoints of ρ and δ are different, a term $\big\langle\Phi_{(3)}, R_2 R_{1'}\big\rangle$ is wedged into the holonomy of ρ which forbids to write the result as H_ρ of something times H_δ of something. Fortunately, letting this operator acting on Ω will eliminate the problem. Applying Theorem 9.2 (ii) to the properly contractible δ within the disk L and using that $\Omega \in \mathscr{P}(L)$, we get

$$\begin{aligned}
R_2 R_{1'} \otimes \mathrm{Hol}_\delta^\circ(R_1 \rightharpoonup \Psi \leftharpoonup R_{2'})\Omega &= R_2 R_{1'} \otimes \Omega \cdot \varepsilon_{\mathcal{D}*}(R_1 \rightharpoonup \Psi \leftharpoonup R_{2'}) \\
&= R_2 R_{1'} \otimes \Omega \cdot \big\langle\Psi, R_{2'} R_1\big\rangle.
\end{aligned}$$

Substituting the definition of the R-matrix, we see that $R_2 R_{1'} \otimes R_{2'} R_1 = (\xi_i \otimes x_j) \otimes (\xi_j \otimes x_i)$, therefore

$$R_2 R_{1'}\big\langle\Psi, R_{2'} R_1\big\rangle = \overleftrightarrow{\Psi}.$$

Thus, the wedge term is simply $\big\langle\Phi_{(3)}, \overleftrightarrow{\Psi}\big\rangle$. If we knew that $\overleftrightarrow{\Psi}$ is a central element of $\mathcal{D}(H)$, then the wedge term would be shiftable along ρ until the endpoint and this would prove (10.1). So it remains to show that $\overleftrightarrow{\Psi} \in \mathrm{Center}\,\mathcal{D}(H)$ if and only if $\Psi \in \mathrm{Cocom}\,\mathcal{D}(H)^*$.

Let us introduce the projections $\mathbf{P}$ and its transpose $\mathbf{P}^\mathsf{T}$ by

$$\mathbf{P}\colon\ \mathcal{D}(H) \to \mathcal{D}(H), \qquad \mathbf{P}(X) := i_{\mathcal{D}(1)} X S_{\mathcal{D}}\big(i_{\mathcal{D}(2)}\big),$$
$$\mathbf{P}^\mathsf{T}\colon\ \mathcal{D}(H)^* \to \mathcal{D}(H)^*, \qquad \mathbf{P}^\mathsf{T}(\Phi) := i_{\mathcal{D}(1)} \rightharpoonup \Phi \leftharpoonup S_{\mathcal{D}}\big(i_{\mathcal{D}(2)}\big),$$

where $i_{\mathcal{D}} := \iota \otimes i$ is the Haar integral of $\mathcal{D}(H)$. $\mathbf{P}$ projects onto $\operatorname{Center}(\mathcal{D}(H))$ and $\mathbf{P}^\mathsf{T}$ onto $\operatorname{Cocom}\mathcal{D}(H)^*$. Using the identity $R\Delta_{\mathcal{D}}(X) = \Delta^{\mathrm{op}}(X)R$ of R-matrices, we obtain

$$
\begin{aligned}
\mathbf{P}(R_2 R_{1'}) \otimes R_{2'} R_1 &= i_{\mathcal{D}(1)} R_2 R_{1'} S_{\mathcal{D}}\big(i_{\mathcal{D}(2)}\big) \otimes R_{2'} R_1 \\
&= i_{\mathcal{D}(1)} R_2 R_{1'} S_{\mathcal{D}}\big(i_{\mathcal{D}(4)}\big) \otimes R_{2'} S_{\mathcal{D}}\big(i_{\mathcal{D}(3)}\big) i_{\mathcal{D}(2)} R_1 \\
&= R_2 i_{\mathcal{D}(2)} R_{1'} S_{\mathcal{D}}\big(i_{\mathcal{D}(4)}\big) \otimes R_{2'} S_{\mathcal{D}}\big(i_{\mathcal{D}(3)}\big) R_1 i_{\mathcal{D}(1)} \\
&= R_2 i_{\mathcal{D}(2)} S_{\mathcal{D}}\big(i_{\mathcal{D}(3)}\big) R_{1'} \otimes S_{\mathcal{D}}\big(i_{\mathcal{D}(4)}\big) R_{2'} R_1 i_{\mathcal{D}(1)} \\
&= R_2 R_{1'} \otimes \mathbf{P}(R_{2'} R_1).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\overset{\leftrightarrow}{\Psi} \in \operatorname{Center}\mathcal{D}(H) \quad &\Leftrightarrow \quad \mathbf{P}(\overset{\leftrightarrow}{\Psi}) = \overset{\leftrightarrow}{\Psi} \\
&\Leftrightarrow \quad \mathbf{P}(R_2 R_{1'})\langle\Psi, R_{2'} R_1\rangle = R_2 R_{1'}\langle\Psi, R_{2'} R_1\rangle \\
&\Leftrightarrow \quad R_2 R_{1'}\langle\mathbf{P}^T(\Psi), R_{2'} R_1\rangle = R_2 R_{1'}\langle\Psi, R_{2'} R_1\rangle \\
&\Leftrightarrow \quad \mathbf{P}^\mathsf{T}(\Psi) = \Psi.
\end{aligned}
$$

So $\overset{\leftrightarrow}{\Psi}$ is central if and only if Ψ is tracial and the theorem is proven. $\blacksquare$

Let D_r be fixed irreducible unitary representations of $\mathcal{D}(H)$ for each isomorphism class r of irreps. Then the set of matrix elements $\big\{D_r^{i,j}\big\}$ is a basis of $\mathcal{D}(H)^*$. For a proper left ribbon $\rho\colon s(a) \to s(b)$ between disjoint sites, the (7.9) and (7.10) imply that the vectors

$$\Lambda_r^{i,j} := \operatorname{Hol}_\rho^\circ\big(D_r^{i,j}\big)\Omega \in \mathscr{H}$$

are irreducible $\mathcal{D}(H)$-multiplets in the sense of the equations

$$
\begin{aligned}
\mathbf{D}_c(X)\Lambda_r^{i,k} &= \Lambda_r^{i,k} \cdot \varepsilon_{\mathcal{D}}(X) \qquad \text{if} \quad c \ne a, b, \\
\mathbf{D}_b(X)\Lambda_r^{i,k} &= \Lambda_r^{i,j} \cdot D_r^{j,k}(X), \\
\mathbf{D}_a(X)\Lambda_r^{i,k} &= D_r^{i,j}(S_{\mathcal{D}}(X)) \cdot \Lambda_r^{j,k}
\end{aligned}
$$

provided $\Omega \in \mathscr{P}$. Notice that $\varepsilon_{\mathcal{D}} = D_0$ if $r = 0$ designates the trivial representation and $D_r^\mathsf{T} \circ S_{\mathcal{D}}$ is, up to unitary equivalence, some $D_{\bar{r}}$ if $r \mapsto \bar{r}$ designates charge conjugation.

Let γ be a closed proper ribbon winding around $s(b)$ but not $s(a)$ as in Theorem 10.1. Let Ψ_r be the unique cocommutative element for which $\overset{\leftrightarrow}{\Psi}_r$ is the minimal central idempotent $e_r = e_r^{i,i}$, where $\big\{e_r^{i,j}\big\}$ is the dual basis of $\big\{D_r^{i,j}\big\}$. Then

$$\operatorname{Hol}_\gamma^\circ(\Psi_q)\Lambda_r^{i,j} = \operatorname{Hol}_\rho^\circ\big(e_q \rightharpoonup D_r^{i,j}\big)\Omega = \Lambda_q^{i,j} \cdot \delta_{q,r}.$$

In this sense, the closed ribbon operators $\operatorname{Hol}_\gamma^\circ(\Psi)$ are able to detect the charge r inside the contraction disk of γ. This happens in spite of the fact that for whatever closed γ for which the ribbon-homotopy class of ρ contains a ρ' disjoint from γ the $\operatorname{Hol}_\gamma^\circ(\Psi)$ detects only the trivial charge on $\Lambda_r^{i,j}$. One says that the charge of $\Lambda_r^{i,j}$ can be localized in semi-infinite strings but not in a finite region. This kind of charges are characteristic of massive gauge theories [2]. This is the first evidence of that the Kitaev model is like a gauge theory. The more precise analysis should chart the superselection sectors similarly to the Doplicher–Haag–Roberts theory [6, 7, 8, 9] but for string localized charges. In the abelian Kitaev model on a square lattice such a program can be carried out [4, 19] but for arbitrary f.d. C*-Hopf algebra H and for arbitrary infinite (planar) OCPM Σ it is not yet in sight.

Acknowledgments

The author wishes to thank the anonymous referees for their thorough reading of the manuscript and for their precious comments which improved considerably the final version of this paper.

References

[1] Bombin H., Martin-Delgado M., A family of non-Abelian Kitaev models on a lattice: topological condensation and confinement, *Phys. Rev. B* **78** (2008), 115421, 28 pages, arXiv:0712.0190.

[2] Buchholz D., Fredenhagen K., Locality and the structure of particle states, *Comm. Math. Phys.* **84** (1982), 1–54.

[3] Buerschaper O., Mombelli J.M., Christandl M., Aguado M., A hierarchy of topological tensor network states, *J. Math. Phys.* **54** (2013), 012201, 46 pages, arXiv:1007.5283.

[4] Cha M., Naaijkens P., Nachtergaele B., On the stability of charges in infinite quantum spin systems, *Comm. Math. Phys.* **373** (2020), 219–264, arXiv:1804.03203.

[5] Cowtan A., Majid S., Quantum double aspects of surface code models, *J. Math. Phys.* **63** (2022), 042202, 49 pages, arXiv:2107.04411.

[6] Doplicher S., Haag R., Roberts J.E., Fields, observables and gauge transformations. I, *Comm. Math. Phys.* **13** (1969), 1–23.

[7] Doplicher S., Haag R., Roberts J.E., Fields, observables and gauge transformations. II, *Comm. Math. Phys.* **15** (1969), 173–200.

[8] Doplicher S., Haag R., Roberts J.E., Local observables and particle statistics. I, *Comm. Math. Phys.* **23** (1971), 199–230.

[9] Doplicher S., Haag R., Roberts J.E., Local observables and particle statistics. II, *Comm. Math. Phys.* **35** (1974), 49–85.

[10] Hirmer A.K., Meusburger C., Categorical generalizations of quantum double models, arXiv:2306.05950.

[11] Jones G., Singerman D., Maps, hypermaps and triangle groups, in The Grothendieck Theory of Dessins D'enfants (Luminy, 1993), *London Math. Soc. Lecture Note Ser.*, Vol. 200, Cambridge University Press, Cambridge, 1994, 115–145.

[12] Kitaev A.Yu., Fault-tolerant quantum computation by anyons, *Ann. Physics* **303** (2003), 2–30, arXiv:quant-ph/9707021.

[13] Lando S.K., Zvonkin A.K., Graphs on surfaces and their applications, *Encyclopaedia Math. Sci.*, Vol. 141, Springer, Berlin, 2004.

[14] Larson R.G., Radford D.E., Semisimple cosemisimple Hopf algebras, *Amer. J. Math.* **110** (1988), 187–195.

[15] Meusburger C., Kitaev lattice models as a Hopf algebra gauge theory, *Comm. Math. Phys.* **353** (2017), 413–468, arXiv:1607.01144.

[16] Meusburger C., Wise D.K., Hopf algebra gauge theory on a ribbon graph, *Rev. Math. Phys.* **33** (2021), 2150016, 93 pages, arXiv:1512.03966.

[17] Montgomery S., Hopf algebras and their actions on rings, *CBMS Reg. Conf. Ser. Math.*, Vol. 82, American Mathematical Society, Providence, RI, 1993.

[18] Naaijkens P., Localized endomorphisms in Kitaev's toric code on the plane, *Rev. Math. Phys.* **23** (2011), 347–373, arXiv:1012.3857.

[19] Naaijkens P., Kitaev's quantum double model from a local quantum physics point of view, in Advances in Algebraic Quantum Field Theory, *Math. Phys. Stud.*, Springer, Cham, 2015, 365–395, arXiv:1508.07170.

[20] Sweedler M.E., Hopf algebras, *Math. Lecture Note Ser.*, W.A. Benjamin, Inc., New York, 1969.

[21] White A.T., Graphs of groups on surfaces. Interactions and models, *North-Holland Math. Stud.*, Vol. 188, North-Holland Publishing Co., Amsterdam, 2001.

[22] Yan B., Chen P., Cui S.X., Ribbon operators in the generalized Kitaev quantum double model based on Hopf algebras, *J. Phys. A* **55** (2022), 185201, 34 pages, arXiv:2105.08202.