A Sound and Complete Equational Theory for 3-Qubit Toffoli-Hadamard Circuits

Matthew Amy
Simon Fraser University
Burnaby, Canada
matt_amy@sfu.ca

Neil J. Ross
Dalhousie University
Halifax, Canada
neil.jr.ross@dal.ca

Scott Wesley
Dalhousie University
Halifax, Canada
scott.wesley@dal.ca

We give a sound and complete equational theory for 3-qubit quantum circuits over the Toffoli-Hadamard gate set $\{X, CX, CCX, H\}$. That is, we introduce a collection of true equations among Toffoli-Hadamard circuits on three qubits that is sufficient to derive any other true equation between such circuits. To obtain this equational theory, we first consider circuits over the Toffoli-K gate set $\{X, CX, CCX, K\}$, where $K = H \otimes H$. The Toffoli-Hadamard and Toffoli-K gate sets appear similar, but they are crucially different on exactly three qubits. Indeed, in this case, the former generates an infinite group of operators, while the latter generates the finite group of automorphisms of the well-known E_8 lattice. We take advantage of this fact, and of the theory of automorphism groups of lattices, to obtain a sound and complete collection of equations for Toffoli-K circuits. We then extend this equational theory to one for Toffoli-Hadamard circuits by leveraging prior work of Li *et al.* on Toffoli-Hadamard operators.

1 Introduction

The *Toffoli-Hadamard* gate set is obtained by extending the classical reversible gate set $\{X, CX, CCX\}$ with the Hadamard gate H. The addition of the Hadamard gate promotes the gate set $\{X, CX, CCX\}$ from one that is universal for classical reversible computation to one that is universal for quantum computation [1, 23]. Because the Hadamard gate can introduce phases of (-1) and produce superpositions, one can think of the addition of the Hadamard gate as a simple way to augment classical reversible computation with these typically quantum features. In turn, this motivates the study of Toffoli-Hadamard circuits [1, 4, 3, 10, 18, 23, 25, 26].

In recent years, an important effort has been made to understand quantum circuits equationally. If G is a set of quantum gates, an *equational theory* for G is given by a set of *equations* (or *relations*) among the circuits over G. The equational theory is *sound* if it only equates circuits that correspond to the same operator, and *complete* if it always equates circuits that correspond to the same operator. Equational theories can be used to optimize and verify quantum circuits in practice, but, more fundamentally, they can illuminate the mathematical structure underlying the gate set G. Sound and complete equational theories have been found for several important gate sets [2, 7, 8, 11, 19, 22].

In this paper, we give a sound and complete equational theory for 3-qubit Toffoli-Hadamard circuits. A presentation for the group of Toffoli-Hadamard operators was given in [18], but the presentation uses 1-, 2-, and 4-level operators as generators. While these operators can be represented by Toffoli-Hadamard circuits, this leads to an unnatural presentation, from the perspective of quantum circuits. What is more, the presentation of [18] contains over 2000 relations, even when restricted to 3-qubit operators. Many of these relations can be presented concisely as relation schemas in the language of operators, but these relations need to be expanded to be stated in the language of circuits. In contrast, our presentation contains only 65 relations, most of which are natural from the perspective of quantum circuits.

Submitted to:
QPL 2024

To obtain our presentation, we first consider circuits over the Toffoli-K gate set $\{X, CX, CCX, K\}$, where $K = H \otimes H$. The Toffoli-Hadamard and Toffoli-K gate sets appear similar, but they are crucially different on exactly three qubits. Indeed, in this case, the former generates an infinite group of operators, while the latter generates the finite group of automorphisms of the well-known E_8 lattice. The correspondence between 3-qubit Toffoli-K circuits and the automorphisms of the E_8 lattice was previously known (see [14, 21]). We take advantage of this correspondence, and of the theory of automorphism groups of lattices, to obtain a sound and complete collection of equations for Toffoli-K circuits. The automorphism group of the E_8 lattice admits a finite Coxeter presentation, which enjoys many geometric and combinatorial properties, and we use *Tietze transformations* to turn the Coxeter presentation of the group of Toffoli-K operators into a concise circuit presentation. We then extend this equational theory to one for Toffoli-Hadamard circuits by building upon [18]. Our paper therefore regards the group of 3-qubit Toffoli-Hadamard circuits as an extension of the automorphism group of the E_8 lattice in order to elucidate its underlying mathematical structure.

The paper is organized as follows. In Section 2, we define three groups of interest. In Section 3, we recall prior results on finite group presentations and we review Tietze transformations. In Section 4, we use the theory of Coxeter groups to obtain a presentation for the group of 3-qubit Toffoli-K circuits using a minimal number of generators. We moreover show that every operator in this group can be represented by a circuit of Toffoli-count at most 120. In Sections 5 and 6, the results of Li *et al.* [18] are used to extend this presentation to a presentation for 3-qubit circuits over the gate set $\{X, CX, CCX, K, CCZ\}$, and then to one for 3-qubit Toffoli-Hadamard circuits. Our approach relies on a large number of derivations and intricate rewriting proofs, which we relegate to several appendices and a supplement [5].

2 Three Groups and Their Generators

Let $\mathbb{Z}$ denote the ring of integers. The *half-integers* $\mathbb{Z}+1/2$ are defined as $\mathbb{Z}+1/2 = \{a+1/2 \mid a \in \mathbb{Z}\}$ and the ring of *dyadic fractions* $\mathbb{D}$ is defined as $\mathbb{D} = \mathbb{Z}[1/2] = \{a/2^k \mid a \in \mathbb{Z} \text{ and } k \in \mathbb{N}\}$. Equivalently, $\mathbb{D}$ is the smallest subring of $\mathbb{Q}$ that contains both $\mathbb{Z}$ and $1/2$. The E_8 *lattice* Γ_8 is the following collection of 8-dimensional vectors,

$$\Gamma_8 = \left\{ x \in \mathbb{Z}^8 \cup (\mathbb{Z}+1/2)^8 \;\middle|\; \sum x_i \equiv 0 \pmod 2 \right\}.$$

In other words, Γ_8 consists of the vectors in $\mathbb{R}^8$ whose components sum to an even integer and are either all integers or all half-integers. The E_8 lattice is well-studied because it enjoys many remarkable properties [12]; in particular, it provides the densest sphere packing in dimension 8 [24].

We now introduce the three groups that will be the focus of this paper. Let R be a ring. For each $n \in \mathbb{N}$, let $\mathrm{GL}(n,R)$ denote the general linear group over R in dimension n and let $\mathrm{O}(n,R)$ denote the orthogonal group over R in dimension n. Define $W(E_8)$ to be the subgroup of $\mathrm{O}(8,\mathbb{D})$ consisting of the elements of $\mathrm{O}(8,\mathbb{D})$ that fix the E_8 lattice. Define $\mathsf{TofH}(n)$ to be the subgroup of $\mathrm{O}(2^n, \mathbb{Z}[1/\sqrt{2}])$ consisting of matrices $M/\sqrt{2}^k$, where M is an integer matrix and $k \in \mathbb{N}$. We will be interested in the groups $W(E_8)$, $\mathrm{O}(8,\mathbb{D})$, and $\mathsf{TofH}(3)$. Note that we have $W(E_8) \leq \mathrm{O}(8,\mathbb{D}) \leq \mathsf{TofH}(3)$.

The above three groups are generated by well-known quantum gates. Let I denote the 2×2 identity matrix and $\otimes$ denote the Kronecker tensor product. Given a dimension 2 matrix M, define $M_0 = M \otimes I \otimes I$, $M_1 = I \otimes M \otimes I$, and $M_2 = I \otimes I \otimes M$. That is, M_j applies operator M to the j-th qubit. Furthermore, define $CM_{j,k}$ to be the operator that sends each standard basis state $|x_0x_1x_2\rangle$ to $(M_k)^{x_j}|x_0x_1x_2\rangle$. That is, $CM_{j,k}$ applies operator M to the k-th qubit whenever the j-th qubit is in the basis state $|1\rangle$. Likewise, define $CCM_{j,k}$ to be the operator that sends each standard basis state $|x_0x_1x_2\rangle$ to $(M_l)^{x_jx_k}|x_0x_1x_2\rangle$ for

$l \in \{1,2,3\} \setminus \{j,k\}$. That is, $CCM_{j,k}$ applies operator M to the l-th qubit whenever the j-th and k-th qubits are both in the basis state $|1\rangle$. The operators $CM_{j,k}$ and $CCM_{j,k}$ denote the usual *controlled-M gate* and *doubly-controlled-M gate*, respectively. Now recall the *Pauli X*, *Pauli Z*, *Hadamard*, and *K* matrices,

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \text{and} \qquad K_{j,k} = H_j \circ H_k,$$

where $(\circ)$ denotes matrix multiplication. Then $CCX_{j,k}$ denotes the *Toffoli* gate. Note that the matrices of the form X_j, $K_{j,k}$, $CX_{j,k}$, $CCX_{j,k}$, and $CCZ_{j,k}$ belong to $\mathrm{O}(8,\mathbb{D})$. It is known that $\{X_j, CX_{j,k}, CCX_{j,k}, K_{j,k}\}$ is a generating set for $W(E_8)$ [12]. Similarly it is known that $\{X_j, CX_{j,k}, CCX_{j,k}, K_{j,k}, CCZ_{j,k}\}$ and $\{X_j, CX_{j,k}, CCX_{j,k}, H_j\}$ are generating sets for $\mathrm{O}(8,\mathbb{D})$ and $\mathsf{TofH}(3)$, respectively [3][1].

3 Presentations and Tietze Transformations

We now recall key results from combinatorial group theory. In particular, we discuss monoid presentations, as well as *Tietze transformations*, which will play an important role in the rest of the paper.

3.1 Presentations

Let Σ be an alphabet (i.e., a set of symbols). Then Σ^* is the *free monoid on* Σ. The elements of Σ^* are the words over Σ, the monoid operation is string concatenation, which we denote by $(\cdot)$, and the identity element in Σ^* is the empty word, which we denote by ε.

If G is a monoid and Q is a quotient of G, then we write $\pi_Q : G \twoheadrightarrow Q$ to denote the canonical projection of G onto Q. Given a subset R of $\Sigma^* \times \Sigma^*$, we write $Q = \langle \Sigma \mid R \rangle$ to denote the largest quotient of $G = \Sigma^*$ such that $\pi_Q(q) = \pi_Q(r)$ for all $(q,r) \in R$. If $M \cong \langle \Sigma \mid R \rangle$, then we say that $\langle \Sigma \mid R \rangle$ is a *presentation of* M and write $q \approx_R$ for each $(q,r) \in R$. The elements of Σ are called *generators* and the elements of R are called *relations*. If, for each $x \in \Sigma$, there exists a $w \in \Sigma^*$ such that $\pi_Q(x \cdot w) = \pi_Q(\varepsilon)$, then M is a group and $\langle \Sigma \mid R \rangle$ is a *monoid presentation for the group M*. In either case, if Σ and R are finite, then $\langle \Sigma \mid R \rangle$ is a *finite presentation*. We distinguish between the presentations $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid R' \rangle$ whenever $R \neq R'$, even if R and R' generate the same quotient.

Certain aspects of presentations can be conveniently expressed in the language of string rewriting. Let Σ be an alphabet and $R \subseteq \Sigma^* \times \Sigma^*$. Fix some $u \in \Sigma^*$ and $v \in \Sigma^*$. If there exists some $(q,r) \in R$ and $s,t \in \Sigma^*$ such that $u = s \cdot q \cdot t$ and $v = s \cdot r \cdot t$, then we write,

$$u \xrightarrow{R} v.$$

If either $u \xrightarrow{R} v$ or $u \xleftarrow{R} v$, then we write $u \overset{R}{\leftrightarrow} v$. We say that *u rewrites to v*, denoted $u \sim_R v$, if either $u = v$ or there exists a finite sequence,

$$u \overset{R}{\leftrightarrow} w_1 \overset{R}{\leftrightarrow} w_2 \overset{R}{\leftrightarrow} \cdots \overset{R}{\leftrightarrow} w_n \overset{R}{\leftrightarrow} v.$$

That is, $(\sim_R)$ is the symmetric, transitive and reflexive closure of $\xrightarrow{R}$. Importantly, $\pi_Q(u) = \pi_Q(v)$ in $Q = \langle \Sigma, R \rangle$ if and only if $u \sim_R v$ [9, Ch. 7]. That is, two words u and v represent the same element in G if and only if the relations in R suffice to rewrite u into v. In this sense, the relations in R define a *complete equational theory* for the monoid Q with respect to the generators Σ. For further information on presentations and on rewriting, the reader is encouraged to consult [17] and [9], respectively.

[1] The generator $CCZ_{j,k}$ is necessary to apply these results to the ancilla-free three-qubit case.

3.2 Tietze Transformations

The transformations, which we state formally below, allow one to add a generator, remove a generator, add a relation, and remove a relation. Let Σ be an alphabet, $R \subseteq \Sigma^* \times \Sigma^*$, and $G = \langle \Sigma \mid R \rangle$ be a monoid.

- **Gen**$(+)$. Let x be a symbol. If $x \notin \Sigma$ and $w \in \Sigma^*$, then $G \cong \langle \Sigma \cup \{x\} \mid R \cup \{x \approx w\} \rangle$.
- **Gen**$(-)$. Let $x \in \Sigma$, $x \approx_R w$, $\Pi = \Sigma \setminus \{x\}$, and $Q = R \setminus \{x \approx w\}$. If $Q \subseteq \Pi^* \times \Pi^*$, then $G \cong \langle \Pi \mid Q \rangle$.
- **Rel**$(+)$. If $q \sim_R r$, then $G \cong \langle \Sigma \mid R \cup \{q \approx r\} \rangle$.
- **Rel**$(-)$. Let $q \approx_R r$ and $Q = R \setminus \{q \approx r\}$. If $q \sim_Q r$, then $G \cong \langle \Sigma \mid Q \rangle$.

The **Gen**$(+)$ rule states that one can add a generator if one also adds a relation defining it in terms of the other generators. The **Gen**$(-)$ rule states that a generator can be removed if it is defined in terms of the other generators, and does not appear in any of the other relations. The **Rel**$(+)$ rule states that if a relation can be derived from the existing ones, then it can be added to the set of relations. Finally, the **Rel**$(-)$ rule conversely states that if a relation can be derived from other relations in the presentation, it is redundant and can be removed.

Tietze transformations are sound and complete for the isomorphism of finite monoid presentations. That is, two presentations $\langle \Sigma \mid R \rangle$ and $\langle \Pi \mid Q \rangle$ specify the same monoid if and only if $\langle \Sigma \mid R \rangle$ can be obtained from $\langle \Pi \mid Q \rangle$ through a finite sequence of Tietze transformations [15, Section 1].

The goal of this paper is to find presentations for groups of quantum operators in which each generator corresponds to a specific quantum gate. More explicitly, given a group G, a generating set Σ, and a *semantic interpretation* $[\![\cdot]\!]_\Sigma : \Sigma \to G$, our goal is to find a set of relations $R \subseteq \Sigma^* \times \Sigma^*$ such that $[\![\cdot]\!]$ induces an isomorphism between $\langle \Sigma \mid R \rangle$ and G. In what follows, we start from a known presentation $\langle \Pi \mid Q \rangle$ over different generators Π with a semantic interpretation $[\![\cdot]\!]_\Pi : \Pi \to G$, and obtain $\langle \Sigma \mid R \rangle$ via a sequence of Tietze transformations. As these Tietze transformations act on the abstract group $\langle \Pi \mid Q \rangle$, one must ensure that the transformations respect the intended interpretation $[\![\cdot]\!]_\Sigma$ of the new generators in Σ, as discussed further in Appendix A.

4 From Coxeter to Circuit Presentations of $W(E_8)$

A *Coxeter group* is a group G which admits a group presentation of the form $\langle r_1, \ldots, r_n \mid (r_j r_k)^{N_{j,k}} \approx \varepsilon \rangle$, where N is an $n \times n$ matrix over $\mathbb{N} \cup \{\infty\}$ such that $N_{j,j} = 1$ and $N_{j,k} > 1$ for all $j \neq k$ [16]. The matrix N is known as the *Coxeter matrix* of G. Note that since $\pi_G(r_j \cdot r_j) = \varepsilon$ for each r_j, then every Coxeter presentation is automatically a monoid presentation for a group. Coxeter groups are an abstraction of reflection groups and, in particular, for every finite Coxeter group G, there is a faithful group representation $G \to \mathrm{O}(n)$ that maps each r_j to a reflection in $\mathbb{R}^n$ [16]. Recall that a *Householder transformation* is a reflection about the hyperplane normal to some vector $\alpha \in \mathbb{R}^n$ defined by $v \mapsto v - 2\frac{\langle v,\alpha \rangle}{\langle \alpha,\alpha \rangle}\alpha$ [16]. If r is the reflection about the hyperplane normal to $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{O}(n)$, then $M \circ r \circ M^{-1}$ is the reflection about the hyperplane normal to $M\alpha$ [16, Prop. 1.2]. As a special case, v and $-v$ define the same reflection.

The goal of this section is to construct a presentation for the Weyl group of the E_8 lattice in terms of Toffoli-K gates. Recall that the Weyl group for any lattice $L \subseteq \mathbb{R}^n$ is the finite reflection group generated by reflections about the roots of L (see [16, Sec. 2.9]). That is, given a root system Φ for L, the group $W(L)$ is generated by $\{r_\alpha : \alpha \in \Phi\}$ where r_α is the reflection through the hyperplane normal to α. Consequently, $W(E_8)$ is a Coxeter group. A root system and the corresponding Coxeter matrix for $W(E_8)$ are given in Figure 1.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1/2 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1/2 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & -1/2 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & -1/2 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1/2 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1/2 \end{bmatrix}$$

(a) E_8 Root System.

$$\begin{bmatrix} 1 & 3 & 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & 1 & 3 & 2 & 2 & 2 & 2 & 2 \\ 2 & 3 & 1 & 3 & 2 & 2 & 2 & 2 \\ 2 & 2 & 3 & 1 & 3 & 2 & 2 & 2 \\ 2 & 2 & 2 & 3 & 1 & 3 & 3 & 2 \\ 2 & 2 & 2 & 2 & 3 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 3 & 2 & 1 & 3 \\ 2 & 2 & 2 & 2 & 2 & 2 & 3 & 1 \end{bmatrix}$$

(b) $W(E_8)$ Coxeter Matrix.

Figure 1: The root system and Coxeter matrix for $W(E_8)$. Note that the root system consists of 8 vectors and are presented as the columns of an 8×8 matrix.

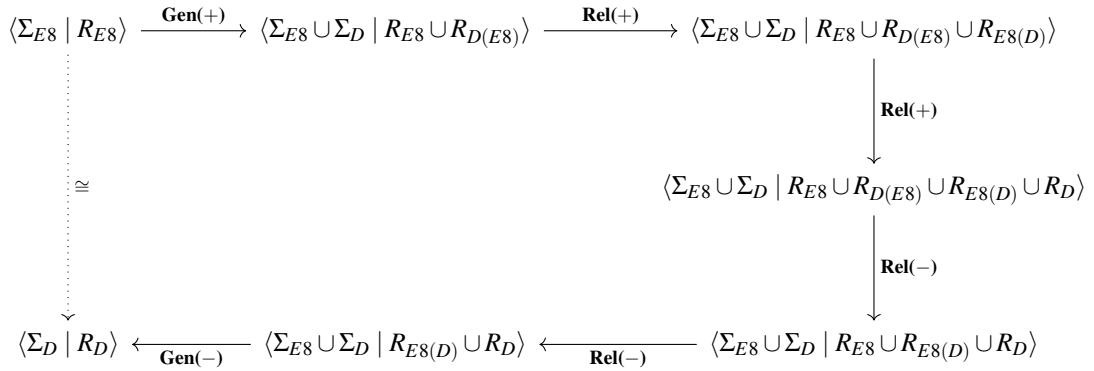

Figure 2: A diagrammatic summary of the Tietze transformations used to obtain a presentation for $W(E_8)$. Note that in this diagram Σ_{E8} denotes the Coxeter generators, Σ_D denotes the dyadic Toffoli-Hadamard gates, $R_{D(E8)}$ expresses the Σ_D in terms of Σ_{E8}, and $R_{E8(D)}$ expresses Σ_{E8} in terms of Σ_D.

To obtain a presentation in terms of Toffoli-K gates, we begin with the Coxeter presentation of $W(E_8)$. The desired presentation is then obtained through a sequence of Tietze transformations. First, the **Gen**$(+)$ rule is used to introduce the dyadic Toffoli-K gates with their intended semantics. Second, the **Rel**$(+)$ rule is used to rewrite the Coxeter generators in terms of Toffoli-K gates (call these relations $R_{E8(D)}$). Third, the **Rel**$(+)$ rule is used to introduce well-known relations satisfied by the Toffoli-K gates (see, e.g., [18, 19]). Given these new relations, the **Rel**$(-)$ rule is used to eliminate all defining relations for the Toffoli-K gates. In a similar fashion, the **Rel**$(-)$ rule is also used to eliminate the Coxeter relations of $W(E_8)$. At this point, the Coxeter generators only appear in $R_{E8(D)}$, and can be eliminated using the **Gen**$(-)$ rule. What remains is a presentation of $W(E_8)$ in terms of Toffoli-K gates. All steps of this proof are summarized in Figure 2.

Each step of this proof requires numerous applications of the corresponding Tietze transformation. To establish that each **Gen**$(+)$ and **Rel**$(+)$ transformation holds, an equation of 8×8 matrices must be validated. To establish that each **Rel**$(-)$ transformation holds, a derivational proof must be validated. In both cases, the proof obligation is computational in nature. The validity of our Tietze transformations have been machine-verified by the software package TIETZE[2].

[2]Available at: https://github.com/meamy/tietze.

4.1 Introducing the Toffoli-K Gates

The generators of $W(E_8)$ can be written as follows.

$$r_1 = X_0 \circ X_1 \circ CCX_{0,1} \circ X_1 \circ X_0 \qquad r_2 = X_0 \circ CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1} \circ X_0$$
$$r_3 = X_0 \circ CCX_{0,1} \circ X_0 \qquad r_4 = CX_{0,1} \circ CX_{0,2} \circ CCX_{1,2} \circ CX_{0,2} \circ CX_{0,1}$$
$$r_5 = X_1 \circ CCX_{0,1} \circ X_1 \qquad r_6 = CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1}$$
$$r_7 = CZ_{0,1} \circ CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1} \circ CZ_{0,1} \qquad r_8 = K_{1,2} \circ X_1 \circ X_2 \circ CZ_{0,2} \circ CCX_{1,2} \circ CZ_{0,2} \circ X_2 \circ X_1 \circ K_{1,2}$$

These equations can be derived from the geometry of $\mathbb{R}^8$. First, note that $CCX_{0,1}$ is a reflection about the hyperplane normal to $|\widehat{b}\rangle = |1\rangle \otimes |1\rangle \otimes |-\rangle$ where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Then for each generator r_j with normal vector $|b_j\rangle$, it suffices to find an element $M \in W(E_8)$ such that $M|\widehat{b}\rangle = |b_j\rangle$. The corresponding circuit would be $M \circ CCX_{1,2} \circ M^{-1}$. As an example of this technique, consider the Coxeter generator r_3 defined by the normal vector $|b_3\rangle = |0\rangle \otimes |1\rangle \otimes |-\rangle$. Since $(X_0)|\widehat{b}\rangle = |b_3\rangle$ with X_0 self-inverse, then $r_3 = X_0 \circ CCX_{0,1} \circ X_0$. The remaining cases are established in Appendix C.

Next, the Toffoli-K gates are introduced. For simplicity of presentation, we first introduce the swap matrices $\sigma_{j,k} = CX_{j,k} \circ CX_{k,j} \circ CX_{j,k}$ where $\sigma_{j,k}$ permutes the j-th qubit with the k-th qubit. Recall that the Toffoli-K gates correspond to the following matrices[3]:

$$\Sigma_D := \left\{ X_j, CX_{k,l}, CCX_{j,k}, Z_j, CZ_{j,k}, K_{j,j+1}, \sigma_{j,k} \mid j,k,l \in \{0,1,2\}, j < k, j \neq l \neq k \right\}.$$

It turns out that all Toffoli-K gates are generated by X_0, $CX_{1,0}$, $CCX_{1,2}$, and $K_{1,2}$. To derive these primitive gates, it helps to first derive several diagonal matrices over (± 1). These are then used to derive the $CCX_{0,1}$ and X_0 gates. From this, the swap matrices can be derived, after which, it is relatively straightforward to construct the X_0, $K_{1,2}$, and $CX_{0,1}$ gates. This yields four words w_X, w_K, w_{CX}, and w_{CCX}, such that $[\![w_X]\!]^*_{E8} = X_0$, $[\![w_{CX}]\!]^*_{E8} = CX_{1,0}$, $[\![w_{CCX}]\!]^*_{E8} = CCX_{1,2}$, and $[\![w_K]\!]^*_{E8} = K_{1,2}$, as outlined in Appendix D.

The remaining Toffoli-K gates are derived in terms of $K_{1,2}$, $CCX_{1,2}$, X_0, and $CX_{0,1}$. To simplify this process, we note that once a gate has been derived, it may then be used to derive other gates. This is analogous to how the generator X_0 appears in the defining relation for $CX_{0,1}$. Given a set of defining relations, if the dependencies between the generators defined by the relations form an acyclic digraph, then the defining relations arise from a valid sequence of **Gen**$(+)$ transformations (see Appendix B.2). Likewise, the derived generators can be eliminated by a valid sequence of Tietze transformation. The defining relations for the remaining 19 gates are found in Appendix B.2. Since the dependencies among these generators are acyclic, then they must arise from 19 valid applications of the **Gen**$(+)$ rules. Let $R_{D(E8)}$ denote all 23 relations. Then $W(E_8) \cong \langle \Sigma_{E8} \cup \Sigma_D \mid R_{E8} \cup R_{D(E8)} \rangle$.

4.2 Deriving the $W(E_8)$ Coxeter Generators

Recall the circuit definitions for the $W(E_8)$ generators from Section 4.1. Let $R_{E8(D)}$ denote the set of corresponding relations. For example, the relation corresponding to r_3 is $r_3 \approx X_0 \cdot CCX_{0,1} \cdot X_0$. Since these relations hold by definition, then they may be introduced via 8 applications of **Rel**$(+)$ and consequently $W(E_8) \cong \langle \Sigma_{E8} \cup \Sigma_D \mid R_{E8} \cup R_{D(E8)} \cup R_{E8(D)} \rangle$.

4.3 Elimination of the Coxeter Generators

In this section, the relations in R_{E8} and $R_{D(E8)}$ are eliminated. To do this, some additional (Σ_D)-relations are required. For the remainder of this section, let $M_{(x_0,x_1,\ldots,x_k)}$ denote a gate M applied to the qubits x_0

[3]For simplicity, we assume that all K gates are applied to adjacent qubits. This is sufficient, since $K_{0,2} = K_{0,1} \circ K_{1,2}$.

$$\begin{aligned}
\mathbf{r_1}\cdot r_1 &\to X_0\cdot X_1\cdot CCX_{0,1}\cdot X_1\cdot X_0\cdot \mathbf{r_1}\\
&\to X_0\cdot X_1\cdot CCX_{0,1}\cdot X_1\cdot \mathbf{X_0}\cdot \mathbf{X_0}\cdot X_1\cdot CCX_{0,1}\cdot X_1\cdot X_0\\
&\to X_0\cdot X_1\cdot CCX_{0,1}\cdot \mathbf{X_1}\cdot \mathbf{X_1}\cdot CCX_{0,1}\cdot X_1\cdot X_0\\
&\to X_0\cdot X_1\cdot \mathbf{CCX_{0,1}}\cdot \mathbf{CCX_{0,1}}\cdot X_1\cdot X_0\\
&\to X_0\cdot \mathbf{X_1}\cdot \mathbf{X_1}\cdot X_0\\
&\to \mathbf{X_0}\cdot \mathbf{X_0}\\
&\to \varepsilon
\end{aligned}$$

(a) Deriving ε from $r_1\cdot r_1$.

$$\begin{aligned}
\mathbf{CX_{0,1}}\cdot X_1 &\to \sigma_{0,1}\cdot CX_{1,0}\cdot \sigma_{0,1}\cdot \mathbf{X_1}\\
&\to \sigma_{0,1}\cdot CX_{1,0}\cdot \sigma_{\mathbf{0,1}}\cdot \sigma_{\mathbf{0,1}}\cdot X_0\sigma_{0,1}\\
&\to \sigma_{0,1}\cdot \mathbf{CX_{1,0}}\cdot \mathbf{X_0}\cdot \sigma_{0,1}\\
&\to \sigma_{0,1}\cdot X_0\cdot CX_{1,0}\sigma_{0,1}\\
&\to \sigma_{\mathbf{0,1}}\cdot \mathbf{X_0}\cdot \sigma_{\mathbf{0,1}}\cdot \sigma_{0,1}\cdot CX_{1,0}\cdot \sigma_{0,1}\\
&\to X_1\cdot \sigma_{\mathbf{0,1}}\cdot \mathbf{CX_{1,0}}\cdot \sigma_{\mathbf{0,1}}\\
&\to X_1\cdot CX_{0,1}
\end{aligned}$$

(b) Deriving $X_1\cdot CX_{0,1}$ from $CX_{0,1}\cdot X_1$.

Figure 3: Examples of derivations proofs which appear in the proof that $W(E_8)\cong\langle\Sigma_D\mid R_0\rangle$.

through to x_k. For example, if M is a doubly-controlled X gate, then $M_{(1,2,0)}$ would correspond to $CCX_{1,2}$. Using this notation, we introduce the following families of relations, denoted R_D.

- **Bifunctoriality**. $M_S\cdot N_T = N_T\cdot M_S$, for all $M_S,N_T\in\Sigma_D$ with $S\cap T=\varnothing$.
- **Symmetry**. $\sigma_{i,j}\cdot M_S\cdot\sigma_{i,j}=M_{\sigma_{i,j}(S)}$, for all $M_S\in\Sigma_D$ and integers $0\leq i<j\leq 3$.
- **Order**. $M\cdot M=\varepsilon$, for all $M\in\Sigma_D$.
- **Commutators**. $M_S\cdot N_T=N_T\cdot w$, for all $M_S,N_T\in\Sigma_D$ with $S\cap T\neq\varnothing$ and $w\in(\Sigma_D)^*$ minimal.

The relation $\sigma_{1,2}\approx CZ_{1,2}\cdot K_{1,2}\cdot CZ_{1,2}\cdot K_{1,2}\cdot CZ_{1,2}\cdot K_{1,2}$ from [7] is also included for simplicity. From these relations, all elements of $R_{D(E8)}$ and R_{E8} can be derived. Since all elements of Σ_{E8} are self-inverse, then it suffices to consider only the upper half of the Coxeter matrix for $W(E_8)$. As an example, consider the relation $r_1\cdot r_1\approx\varepsilon$ in R_{E8}. The proof proceeds as in Figure 3a. Then ε can be derived from $r_1\cdot r_1$ using the relations in R_D. Similar methods can be used to eliminate the remaining (R_{E8})-relations. All derivations, for both $R_{D(E8)}$ and R_{E8} can be found in the supplement to this paper [5].

Theorem 4.1. $W(E_8)\cong\langle\Sigma_D\mid R_D\rangle$

It follows immediately from Theorem 4.1 that given any circuit C over Σ_D, there exists a minimal word w over the alphabet $\{r_1,r_2,\ldots,r_8\}$ such $[\![C]\!]^*_D=[\![w]\!]^*_{E8}$. Then by the decompositions of Section 4.1, there exists a circuit C' over Σ_D such that $[\![w]\!]^*_{E8}=[\![C']\!]^*_D$ such that C' contains exactly $|w|$ Toffoli gates. By [16, Thm. 1.8], every minimal word in $W(E_8)$ has length at most n, where n is the cardinality of the positive root system associated with $W(E_8)$. By [16], the positive root system associated with $W(E_8)$ has cardinality 120. Therefore, C' contains at most 120 Toffoli gates. This provides an upper-bound on the Toffoli count for circuits over Σ_D, which can be thought of as a measure of computational complexity for these three-qubit circuits.

Corollary 4.2. *If $C\in\Sigma_D^*$, then there exists $C'\in\Sigma_D$ with Toffoli count at most* 120 *such that $[\![C]\!]^*_D=[\![C']\!]^*_D$.*

4.4 A Reduced Set of Relations for $W(E_8)$

The relations R_D from Section 4.3 are far from minimal. For example, the family of commutator relations contains all relations of the form $CX_{j,k}\cdot X_k\approx X_k\cdot CX_{j,k}$. However, given all symmetry relations, it suffices to include only $CX_{1,0}\cdot X_0\approx X_0\cdot CX_{1,0}$. The remaining commutator relations can be derived, as illustrated in Figure 3b. Furthermore, many of the relations in R_D do not appear in any derivations of the supplement. For example, the relation $CX_{2,0}\cdot X_0\approx X_0\cdot CX_{2,0}$ does not appear, and therefore Theorem 4.1 holds with respect to the relation set $R_D\setminus\{CX_{2,0}\cdot X_0\approx X_0\cdot CX_{2,0}\}$. Using both techniques, a new relation set R_0 is obtained, as illustrated in Figure 4. All derivations can be found in the supplement to this paper [5].

Corollary 4.3. $W(E_8)\cong\langle\Sigma_D\mid R_0\rangle$

$$CZ_{0,1} \approx K_{1,2} \cdot CX_{0,1} \cdot K_{1,2} \quad (1)$$
$$X_1 \approx CX_{0,1} \cdot X_0 \cdot CX_{0,1} \cdot X_0 \quad (2)$$
$$Z_0 \approx CZ_{0,1} \cdot CX_{0,1} \cdot CZ_{0,1} \cdot CX_{0,1} \quad (3)$$
$$Z_1 \approx K_{1,2} \cdot X_1 \cdot K_{1,2} \quad (4)$$
$$CX_{2,0} \approx X_1 \cdot CCX_{1,2} \cdot X_1 \cdot CCX_{1,2} \quad (5)$$
$$CX_{2,1} \approx CX_{2,0} \cdot CX_{0,1} \cdot CX_{2,0} \cdot CX_{0,1} \quad (6)$$
$$CX_{1,2} \approx K_{1,2} \cdot CX_{2,1} \cdot K_{1,2} \quad (7)$$
$$\sigma_{1,2} \approx CX_{1,2} \cdot CX_{2,1} \cdot CX_{1,2} \quad (8)$$
$$CX_{0,2} \approx \sigma_{1,2} \cdot CX_{0,1} \cdot \sigma_{1,2} \quad (9)$$
$$\sigma_{0,2} \approx CX_{0,2} \cdot CX_{2,0} \cdot CX_{0,2} \quad (10)$$
$$K_{0,1} \approx \sigma_{0,2} \cdot K_{1,2} \cdot \sigma_{0,2} \quad (11)$$
$$CX_{1,0} \approx K_{0,1} \cdot CX_{0,1} \cdot K_{0,1} \quad (12)$$
$$\sigma_{0,1} \approx CX_{0,1} \cdot CX_{1,0} \cdot CX_{0,1} \quad (13)$$
$$CCX_{0,2} \approx \sigma_{0,1} \cdot CCX_{1,2} \cdot \sigma_{0,1} \quad (14)$$
$$X_2 \approx \sigma_{0,2} \cdot X_0 \cdot \sigma_{0,2} \quad (15)$$
$$Z_2 \approx \sigma_{0,2} \cdot Z_0 \cdot \sigma_{0,2} \quad (16)$$
$$CCX_{0,1} \approx K_{1,2} \cdot CCX_{0,2} \cdot K_{1,2} \quad (17)$$
$$CZ_{0,2} \approx \sigma_{1,2} \cdot CZ_{0,1} \cdot \sigma_{1,2} \quad (18)$$
$$CZ_{1,2} \approx \sigma_{0,1} \cdot CZ_{0,2} \cdot \sigma_{0,1} \quad (19)$$
$$\sigma_{1,2} \approx CZ_{1,2} \cdot K_{1,2} \cdot CZ_{1,2} \cdot K_{1,2} \cdot CZ_{1,2} \cdot K_{1,2} \quad (20)$$
$$X_0 \cdot X_0 \approx \varepsilon \quad (21)$$
$$CX_{0,1} \cdot CX_{0,1} \approx \varepsilon \quad (22)$$
$$K_{1,2} \cdot K_{1,2} \approx \varepsilon \quad (23)$$
$$CCX_{1,2} \cdot CCX_{1,2} \approx \varepsilon \quad (24)$$
$$K_{0,1} \cdot K_{0,1} \approx \varepsilon \quad (25)$$
$$CX_{1,2} \cdot X_0 \approx X_0 \cdot CX_{1,2} \quad (26)$$
$$X_0 \cdot K_{1,2} \approx K_{1,2} \cdot X_0 \quad (27)$$
$$X_1 \approx \sigma_{0,1} \cdot X_0 \cdot \sigma_{0,1} \quad (28)$$
$$CX_{2,0} \approx \sigma_{0,2} \cdot CX_{0,2} \cdot \sigma_{0,2} \quad (29)$$
$$CX_{1,2} \approx \sigma_{0,1} \cdot CX_{0,2} \cdot \sigma_{0,1} \quad (30)$$
$$CX_{2,1} \approx \sigma_{0,1} \cdot CX_{2,0} \cdot \sigma_{0,1} \quad (31)$$
$$CCX_{0,1} \approx \sigma_{0,2} \cdot CCX_{1,2} \cdot \sigma_{0,2} \quad (32)$$
$$CCX_{0,1} \approx \sigma_{1,2} \cdot CCX_{0,2} \cdot \sigma_{0,2} \quad (33)$$
$$Z_1 \approx \sigma_{0,1} \cdot Z_0 \cdot \sigma_{0,1} \quad (34)$$
$$K_{0,1} \approx \sigma_{0,1} \cdot K_{0,1} \cdot \sigma_{0,1} \quad (35)$$
$$CCX_{1,2} \cdot CX_{1,0} \approx CX_{1,0} \cdot CCX_{1,2} \quad (36)$$
$$X_0 \cdot CCX_{1,2} \approx CCX_{1,2} \cdot X_0 \quad (37)$$
$$X_0 \cdot CX_{1,0} \approx CX_{1,0} \cdot X_0 \quad (38)$$
$$K_{0,1} \cdot K_{1,2} \approx K_{1,2} \cdot K_{0,1} \quad (39)$$
$$CZ_{0,1} \cdot CZ_{1,2} \approx CZ_{1,2} \cdot CZ_{0,1} \quad (40)$$
$$K_{0,1} \cdot Z_0 \approx X_0 \cdot K_{0,1} \quad (41)$$
$$X_0 \cdot CCX_{0,1} \approx CCX_{0,1} \cdot CX_{1,2} \cdot X_0 \quad (42)$$
$$CX_{0,1} \cdot CZ_{1,2} \approx CZ_{1,2} \cdot CZ_{0,2} \cdot CX_{0,1} \quad (43)$$
$$CX_{1,2} \cdot CCX_{1,2} \approx CCX_{1,2} \cdot CX_{1,0} \cdot CX_{1,2} \quad (44)$$
$$CCX_{1,2} \cdot CX_{0,1} \approx CX_{0,1} \cdot CCX_{0,2} \cdot CCX_{1,2} \cdot CCX_{0,2} \quad (45)$$
$$CCX_{0,1} \cdot CCX_{0,2} \approx CCX_{0,2} \cdot CCX_{0,1} \cdot CCX_{0,2} \cdot CCX_{0,1} \quad (46)$$

Figure 4: Relations for $W(E_8)$, denoted R_0.

4.5 A Minimal Generating Set for $W(E_8)$

Define $\Sigma_0 = \{X_0, CX_{0,1}, CCX_{1,2}, K_{1,2}\}$. From Section 4.1, it is clear that Σ_0 generates $W(E_8)$. In fact, Σ_0 is minimal in the sense that every proper subset of Σ_0 generates a proper subgroup of $W(E_8)$. In other words, no proper subset of Σ_0 generates $W(E_8)$. To show that Σ_0 is a minimal generating set for $W(E_8)$, it suffices to show that for every maximal proper subset Σ' of Σ_0, there exists some 8×8 dyadic matrix M such that M commutes with the elements of Σ' but does not commute with the elements of Σ_0. Intuitively, the subgroup generated by Σ' commutes with M, whereas the subgroup generated by Σ_0 does not commute with M. This claim is proven in Appendix E, and the matrices are constructed.

Theorem 4.4. *Σ_0 is a minimal generating set for $W(E_8)$.*

5 Extending to a Presentation of $\mathrm{O}(8,\mathbb{D})$

Li *et al.* [18] introduced a presentation for $\mathrm{O}(8,\mathbb{D})$ using m-level operators. Let $n > 0$, I be the $n \times n$ identity matrix, and $[m] = \{0, 1, \ldots, m-1\}$. Then given an $m \times m$ matrix M with $m < n$, and a strictly increasing sequence $(a_0, \ldots, a_{m-1})$ over $[m]$, define $M_{[a_0,\ldots,a_{m-1}]}$ to be the $n \times n$ matrix such that:

1. For each pair of elements (j,k) over $[m]$, the component (a_j, a_k) of $M_{[a_0,\ldots,a_{m-1}]}$ is equal to the component (a_j, a_k) of M;
2. For each pair of elements (j,k) over $[n] \setminus \{a_0, a_1, \ldots, a_m\}$, the component (j,k) of $M_{[a_0,a_1,\ldots,a_{m-1}]}$ is equal to the component (a_j, a_k) of I.

$$X_{[a,b]}^2 \approx \varepsilon \quad (47)$$
$$(-1)_{[a]}^2 \approx \varepsilon \quad (48)$$
$$K_{[a,b,c,d]}^2 \approx \varepsilon \quad (49)$$
$$X_{[a,b]} \cdot X_{[c,d]} \approx X_{[c,d]} \cdot X_{[a,b]} \quad (50)$$
$$X_{[a,b]} \cdot (-1)_{[c]} \approx (-1)_{[c]} \cdot X_{[a,b]} \quad (51)$$
$$X_{[a,b]} \cdot K_{[c,d,e,f]} \approx K_{[c,d,e,f]} \cdot X_{[a,b]} \quad (52)$$
$$(-1)_{[a]} \cdot K_{[b,c,d,e]} \approx K_{[b,c,d,e]} \cdot (-1)_{[a]} \quad (53)$$
$$(-1)_{[a]} \cdot (-1)_{[b]} \approx (-1)_{[b]} \cdot (-1)_{[a]} \quad (54)$$
$$K_{[a,b,c,d]} \cdot K_{[e,f,g,h]} \approx K_{[e,f,g,h]} \cdot K_{[a,b,c,d]} \quad (55)$$
$$X_{[a,c]} \cdot X_{[a,b]} \approx X_{[c,b]} \cdot X_{[a,c]} \quad (56)$$
$$X_{[b,c]} \cdot X_{[a,b]} \approx X_{[a,c]} \cdot X_{[b,c]} \quad (57)$$
$$X_{[a,b]} \cdot (-1)_{[a]} \approx (-1)_{[b]} \cdot X_{[a,b]} \quad (58)$$
$$X_{[a,e]} \cdot K_{[a,b,c,d]} \approx K_{[e,b,c,d]} \cdot X_{[a,e]} \quad (59)$$
$$X_{[b,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,e,c,d]} \cdot X_{[b,e]} \quad (60)$$
$$X_{[c,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,e,d]} \cdot X_{[c,e]} \quad (61)$$
$$X_{[d,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,c,e]} \cdot X_{[d,e]} \quad (62)$$
$$X_{[a,b]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,c,d]} \cdot X_{[a,b]} \cdot (-1)_{[b]} \cdot (-1)_{[d]} \quad (63)$$
$$X_{[b,c]} \cdot K_{[a,b,c,d]} \approx (-1)_{[a]} \cdot K_{[a,b,c,d]} \cdot (-1)_{[a]} \cdot K_{[a,b,c,d]} \cdot (-1)_{[a]} \quad (64)$$
$$X_{[c,d]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,c,d]} \cdot X_{[b,d]} \quad (65)$$
$$K_{[a,b,c,d]} \cdot K_{[b,d,e,f]} \approx K_{[b,d,e,f]} \cdot K_{[a,b,c,d]} \quad (66)$$
$$(-1)_{[a]} \cdot (-1)_{[e]} \cdot X_{[a,e]} \cdot \rho_{a,b,c,d,e,f,g,h} \approx \rho_{a,b,c,d,e,f,g,h} \cdot X_{[a,e]} \cdot (-1)_{[e]} \cdot (-1)_{[a]} \quad (67)$$

Figure 5: The relations in $\mathscr{R}_n$ from [18], for all valid choices of $a,b,c,d,e,f,g,h \in \mathbb{Z}$. We write $\rho_{a,b,c,d,e,f,g,h}$ for the substring $K_{[e,f,g,h]} \cdot K_{[a,b,c,d]} \cdot X_{[d,e]} \cdot K_{[a,b,c,d]} \cdot K_{[e,f,g,h]}$.

We say that $M_{[a_0,\ldots,a_{m-1}]}$ is an *m-level operator of type M*. When $n = 8$ for example, $CCX_{0,1} = X_{[6,7]}$, $CCZ_{0,1} = (-1)_{[7]}$, and $K_{[4,5,6,7]}$ is a controlled K-gate. Define the following for $n > 3$.

$$\mathscr{G}_n = \left\{(-1)_{[a]}, X_{[a,b]}, K_{[a,b,c,d]} \mid a,b,c,d \in \mathbb{Z} \text{ and } 0 \leq a < b < c < d < n\right\}$$

It was shown in [18] that $\mathrm{O}(n,\mathbb{D}) \cong \langle \mathscr{G}_n \mid \mathscr{R}_n \rangle$, where $\mathscr{R}_n$ is given in Figure 5. The goal of this section is to construct a sequence of Tietze transformations, starting from $\langle \mathscr{G}_8 \mid \mathscr{R}_8 \rangle$, such that the generators and relations describing the subgroup $W(E_8)$ are replaced by Σ_D and R_0, respectively. This process follows similarly to Section 4. However, one should note that $|\mathscr{R}_8| = 2113$ (see Appendix F.1). Inspection of $\mathscr{R}_n$ reveals that many of these relations are either definitional, or obtained through permutations of indices. For this reason, $\mathscr{R}_n$ is partially reduced before carrying out the aforementioned Tietze transformations. First, the permutations are eliminated via a sequence of **Rel**$(-)$ transformations to obtain $\mathscr{R}_n^1$. Next, some redundant commutator relations are eliminated via a sequence of **Rel**$(-)$ transformations to obtain $\mathscr{R}_n^2$. Finally, the derived generators are eliminated to obtain $\mathscr{R}_n^3$. All proofs can be found in Appendix F.

5.1 Permutation Groups and Reindexing

Let $[n] = \{0,1,\ldots,n-1\}$ and $\mathrm{S}(n)$ denote the group of permutations on $[n]$. For $j,k \in [n]$, let $\tau_{j,k}$ denote the permutation that swaps j and k. For example, $\tau_{0,1}(0) = 1$, $\tau_{0,1}(1) = 0$, and $\tau_{0,1}(2) = 2$. The group $\mathrm{S}(n)$ is a finite reflection group generated by the *transpositions* $\{\tau_{j,j+1} \mid j \in [n]\}$ (see [16]). The *braiding relations*, which state that $\tau_{j,j+1} \circ \tau_{j+1,j+2} \circ \tau_{j,j+1} = \tau_{j+1,j+2} \circ \tau_{j,j+1} \circ \tau_{j+1,j+2}$ for all $j \in [n-2]$, together with the order relations are sound and complete for $\mathrm{S}(n)$ (see [17]). The standard representation of $\mathrm{S}(n)$ as a reflection group sends each $\tau_{j,k}$ to $X_{[j,k]}$. This means that every two-level operator of type X can be decomposed into sequence of transpositions. Intuitively, each $X_{[j,k]}$ acts by permuting the standard basis vectors $|j\rangle$ and $|k\rangle$, which can be achieved through a sequence of transpositions of basis vectors. Clearly, $\mathrm{S}(8) \hookrightarrow W(E_8) \leq O(8,\mathbb{D})$.

Many relations in $\mathscr{R}_n$ are related via permutation of indices. The *formal application of* σ to a word over $\mathscr{G}_n$ is defined inductively as follows.

$$\sigma(\varepsilon) = \varepsilon \qquad \sigma(X_{[a,b]} \cdot w) = X_{[\sigma(a),\sigma(b)]} \cdot \sigma(w)$$
$$\sigma((-1)_{[a]} \cdot w) = (-1)_{[\sigma(a)]} \cdot \sigma(w) \qquad \sigma(K_{[a,b,c,d]} \cdot w) = K_{[\sigma(a),\sigma(b),\sigma(c),\sigma(d)]} \cdot \sigma(w)$$

Note that $\sigma(w)$ may yield m-level operators with invalid indices. For example, $\tau_{1,2}(X_{[1,2]})$ yields $X_{[2,1]}$, which is not a valid two-level operator since $2 > 1$. The permutation σ is a *valid reindexing for* w if all symbols in $\sigma(w)$ are well-formed multi-level operators. If σ is valid for v and w, then σ is valid for $v \cdot w$. Conversely, if σ is valid for w, then σ is valid for all subwords in w. Consider, for example, the word $w = K_{[2,3,4,5]} \cdot K_{[3,5,6,7]}$ which appears on the left-hand side of an instance of Relation (55). Let $\sigma \in \mathrm{S}(8)$ be the cyclic permutation $7 \mapsto 5 \mapsto 3 \mapsto 1 \mapsto 6 \mapsto 4 \mapsto 2 \mapsto 0 \mapsto 7$. Then σ is a valid reindexing for w since $\sigma(w) = K_{[0,1,2,5]} \cdot K_{[1,3,4,5]}$. In Appendix F.2, we show that all valid reindexings are derivable using only the relations in $\mathscr{R}_\sigma = \{$Relations (47), (51), (52), (56), (57), (58), (59), (60), (61) and (62)$\}$.

5.2 Selecting Representative Relations for $\mathrm{O}(n, \mathbb{D})$

As a consequence of Section 5.1, many relations in $\mathscr{R}_n$ can be replaced by representative instances. For example, let r denote instance $(-1)_{[6]} \cdot (-1)_{[7]} \approx (-1)_{[7]} \cdot (-1)_{[6]}$ of Relation (54). Clearly $\sigma = \tau_{0,6} \circ \tau_{1,7}$ is a valid reindexing for r, where $\sigma(r)$ is $(-1)_{[0]} \cdot (-1)_{[1]} \approx (-1)_{[1]} \cdot (-1)_{[0]}$ Then by Appendix F.2, it is possible to derive $\sigma(r)$ from r using $\mathscr{R}_n \setminus \{\sigma(r)\}$. Then $\langle \mathscr{G}_n \mid \mathscr{R}_n \rangle \cong \langle \mathscr{G}_n \mid \mathscr{R}_n \setminus \{\sigma(r)\} \rangle$ by **Rel**$(-)$.

This process can be repeated, until all instances of Relation (54) have been eliminated, except for the representative relation r. In a similar fashion, Relations (48), (49), (53), (55), (63), (64), (65), (66) and (67) can be eliminated, since these relations do not appear in $\mathscr{R}_\sigma$. Then $\mathrm{O}(n, \mathbb{D}) \cong \langle \mathscr{G}_n \mid \mathscr{R}_n^1 \rangle$ where $\mathscr{R}_n^1$ is the set of representative relations (see Appendix F.3).

5.3 Selecting Representative Generators for $\mathrm{O}(n, \mathbb{D})$

Define the new generator set,

$$\mathscr{G}_n^1 = \left\{ X_{[a,b]}, \mid a, b, \in \mathbb{Z} \text{ and } 0 \leq a < b < n \right\} \cup \left\{ K_{[0,1,2,3]} \right\} \cup \left\{ (-1)_{[0]} \right\}.$$

Many of the generators in $\mathscr{G}_n$ are redundant in the sense that they may be constructed using only the generators in $\mathscr{G}_n^1$. This is because $\mathrm{S}(n) \hookrightarrow \mathscr{G}_n^1$, with $\mathscr{G}_n^1 \setminus \mathscr{G}_n$ consisting of valid indexings of either $K_{[0,1,2,3]}$ or $(-1)_{[0]}$. Furthermore, these reindexings follow from relations in $\mathscr{R}_n^1$. As an example, consider the instance $X_{[0,7]} \cdot (-1)_{[0]} \approx (-1)_{[7]} \cdot X_{[0,7]}$ of Relation (58). Using the order relation for $X_{[0,7]}$, the following derivation holds.

$$(-1)_{[7]} \leftarrow (-1)_{[7]} \cdot {X_{[0,7]}}^2 \leftarrow X_{[0,7]} \cdot (-1)_{[0]} \cdot X_{[0,7]}$$

Similarly, the original relation can be obtained from this new relation using the order relation for $X_{[0,7]}$. Then through a **Rel**$(+)$ transformation followed by a **Rel**$(-)$ transformation, the commutator relation $X_{[0,7]} \cdot (-1)_{[0]} \approx (-1)_{[7]} \cdot X_{[0,7]}$ can be replaced by the definitional relation $(-1)_{[7]} = X_{[0,7]} \cdot (-1)_{[0]} \cdot X_{[0,7]}$. This process can be repeated for all instances of Relation (58).

To derive the four-level operators of type K, it suffices to note that the following family of relations are valid with respect to $[\![\cdot]\!]_O^*$.

$$K_{[a,b,c,d]} \approx X_{[0,a]} \cdot X_{[1,b]} \cdot X_{[2,c]} \cdot X_{[3,d]} \cdot K_{[0,1,2,3]} \cdot X_{[3,d]} \cdot X_{[2,c]} \cdot X_{[1,b]} \cdot X_{[0,a]}$$

The cases where $\{a,b,c,d\} \cap \{0,1,2,3\} \neq \varnothing$ can be handled using the techniques of Appendix F.2. These relations are introduced using a sequence of **Rel**$(+)$ relations to obtain a new relation set R. In this relation set, all multi-level operators of type (-1) and K are defined in terms of $(-1)_{[0]}$ and $K_{[0,1,2,3]}$, respectively. As outlined in Appendix B.2, these defining relations can be used to eliminate all generators in $\mathscr{G}_n^1 \setminus \mathscr{R}_n^0$ via a finite sequence of Tietze transformations.

$$X_{[a,a+1]}{}^2 \approx \varepsilon \quad (68)$$
$$(-1)_{[0]}{}^2 \approx \varepsilon \quad (69)$$
$$K_{[0,1,2,3]}{}^2 \approx \varepsilon \quad (70)$$
$$X_{[b,b+1]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[b,b+1]} \quad (71)$$
$$X_{[c,c+1]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[c,c+1]} \quad (72)$$
$$(-1)_{[4]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot (-1)_{[4]} \quad (73)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \approx (-1)_{[4]} \cdot (-1)_{[0]} \quad (74)$$
$$K_{[0,1,2,3]} \cdot K_{[4,5,6,7]} \approx K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \quad (75)$$
$$X_{[a,a+1]} \cdot X_{[a,a+2]} \approx X_{[a+1,a+2]} \cdot X_{[a,a+1]} \quad (76)$$
$$X_{[a+1,b]} \cdot X_{[a,a+1]} \approx X_{[a,b]} \cdot X_{[a+1,b]} \quad (77)$$
$$X_{[0,1]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[0,1]} \cdot (-1)_{[1]} \cdot (-1)_{[3]} \quad (78)$$
$$X_{[1,2]} \cdot K_{[0,1,2,3]} \approx (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \quad (79)$$
$$X_{[2,3]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[1,3]} \quad (80)$$
$$K_{[0,1,2,3]} \cdot K_{[1,3,4,5]} \approx K_{[1,3,4,5]} \cdot K_{[0,1,2,3]} \quad (81)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \cdot X_{[0,4]} \cdot \rho \approx \rho \cdot X_{[0,4]} \cdot (-1)_{[4]} \cdot (-1)_{[0]} \quad (82)$$

Figure 6: The reduced relations in $\mathscr{R}_n^3$, for all valid choices of $a,b,c \in \mathbb{Z}$ where $b > 0$ and $c > 3$. We write ρ for the substring $K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \cdot X_{[3,4]} \cdot K_{[0,1,2,3]} \cdot K_{[4,5,6,7]}$.

The elimination process works as follows. Let $M \in \mathscr{G}_n^1 \setminus \mathscr{G}_n$. Then M appears in some defining relation $M \approx w$. If M appears in some relation $r \in R$, then every instance of M will be replaced by w. For example, Relation (54) will be replaced by the following relation.

$$(-1)_{[0]} \cdot X_{[0,5]} \cdot (-1)_{[0]} \cdot X_{[0,5]} \approx X_{[0,5]} \cdot (-1)_{[0]} \cdot X_{[0,5]} \cdot (-1)_{[0]}$$

We introduce the following abbreviations for simplicity of presentation.

$$(-1)_{[c]} = X_{[0,c]} \cdot (-1)_{[0]} \cdot X_{[0,c]} \qquad K_{[0,1,2,d]} = X_{[3,d]} \cdot K_{[0,1,2,3]} \cdot X_{[3,d]} \qquad K_{[0,1,c,d]} = X_{[2,c]} \cdot K_{[0,1,2,d]} \cdot X_{[2,c]}$$
$$K_{[0,b,c,d]} = X_{[1,b]} \cdot K_{[0,1,c,d]} \cdot X_{[1,b]} \qquad K_{[a,b,c,d]} = X_{[0,a]} \cdot K_{[0,b,c,d]} \cdot X_{[0,a]}$$

Denote this new set of relations $\mathscr{R}_n^2$. Then $\mathrm{O}(n,\mathbb{D}) \cong \langle \mathscr{G}_n^1 \mid \mathscr{R}_n^2 \rangle$.

5.4 Eliminating Redundant Relations

It will now shown that many relations in $\mathscr{R}_n^2$ are redundant. First, the braiding relations and order relations are used according to Appendix F.2 to eliminate all other relations over the two-level operators of type X. This reduced relation set is then used to show that all instances of Relations (51) and (52) can be derived using transpositions in place of swaps. Finally, it is shown that the relations Relations (59), (60), (61) and (62) are entirely redundant. All derivations can be found in Appendix F.4. This new set of relations is denoted $\mathscr{R}_n^3$, and can be found in Figure 6. Then via a sequence of **Rel**$(-)$ transformations, the following presentation is obtained.

Theorem 5.1. $\mathrm{O}(n,\mathbb{D}) \cong \langle \mathscr{G}_n^1 \mid \mathscr{R}_n^3 \rangle$.

5.5 Introducing the $W(E_8)$ Generators

It this section, the circuit generators and relations for $W(E_8)$ are introduced. Since $CCX_{1,2} = X_{[6,7]}$, then without loss of generality, every instance of $X_{[6,7]}$ in $\mathscr{R}_8^3$ can be replaced by $CCX_{1,2}$. Next, the generators X_0 and $CX_{0,1}$ are introduced. This yields the following relations.

$$(r_X) : X_0 \approx X_{[0,4]} \cdot X_{[1,5]} \cdot X_{[2,6]} \cdot X_{[3,7]} \qquad (r_{CX}) : CX_{0,1} \approx X_{[2,6]} \cdot X_{[3,7]}$$

It turns out that the $K_{1,2}$ gate decomposes into a word over X_0 and $K_{[0,1,2,3]}$. This is because $K_{[0,1,2,3]}$ is a $K_{1,2}$ gate which is applied when qubit 0 is in state $|1\rangle$, and $X_0 \circ K_{[0,1,2,3]} \circ X_0$ is a $K_{1,2}$ gate which is applied

$$(-1)_{[0]}{}^2 \approx \varepsilon \quad (83)$$
$$K_{[0,1,2,3]}{}^2 \approx \varepsilon \quad (84)$$
$$X_{[1,2]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[1,2]} \quad (85)$$
$$X_{[2,3]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[2,3]} \quad (86)$$
$$X_{[3,4]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[3,4]} \quad (87)$$
$$X_{[4,5]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[4,5]} \quad (88)$$
$$X_{[5,6]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[5,6]} \quad (89)$$
$$X_{[6,7]} \cdot (-1)_{[0]} \approx (-1)_{[0]} \cdot X_{[6,7]} \quad (90)$$
$$X_{[4,5]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[4,5]} \quad (91)$$
$$X_{[5,6]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[5,6]} \quad (92)$$
$$X_{[6,7]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[6,7]} \quad (93)$$
$$(-1)_{[4]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot (-1)_{[4]} \quad (94)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \approx (-1)_{[4]} \cdot (-1)_{[0]} \quad (95)$$
$$K_{[0,1,2,3]} \cdot K_{[4,5,6,7]} \approx K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \quad (96)$$
$$X_{[0,1]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[0,1]} \cdot (-1)_{[1]} \cdot (-1)_{[3]} \quad (97)$$
$$X_{[1,2]} \cdot K_{[0,1,2,3]} \approx (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \quad (98)$$
$$X_{[2,3]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[1,3]} \quad (99)$$
$$K_{[0,1,2,3]} \cdot K_{[1,3,4,5]} \approx K_{[1,3,4,5]} \cdot K_{[0,1,2,3]} \quad (100)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \cdot X_{[0,4]} \cdot \rho \approx \rho \cdot X_{[0,4]} \cdot (-1)_{[4]} \cdot (-1)_{[0]} \quad (101)$$

Figure 7: The relations in $\mathscr{R}_8^4$, sufficient to extend from $W(E_8)$ to $O(8,\mathbb{D})$.

when qubit 0 is in state $|0\rangle$. Together, these two words compose to a $K_{1,2}$ gate without any controls. This yields the following relation.

$$(r_K) : K_{1,2} \approx K_{[4,5,6,7]} \cdot X_0 \cdot K_{[4,5,6,7]} \cdot X_0$$

The relations r_X, r_{CX}, and r_K can be validated with respect to $[\![\cdot]\!]_O$. These relations do not depend on one-another, so the generators in Σ_0 may be introduced via a sequence of **Gen**(+) transformations, as outlined in Appendix B.2. Likewise, the derived generators in $\Sigma_D \setminus \Sigma_0$ may be introduced via a sequence of **Gen**(+) transformations, as outlined in Appendix B.2. Finally, the relations in R_0 may be introduced, since Section 4.1 established the validity of these relations in $W(E_8)$, which is a subgroup of $O(8,\mathbb{D})$. This sequence of transformations yields $\mathrm{O}(8,\mathbb{D}) \cong \langle \mathscr{G}_8^1 \cup \Sigma_D \mid \mathscr{R}_8^3 \cup R_0 \cup \{r_X, r_{CX}, r_K\} \rangle$.

5.6 Eliminating the Multi-Level Operators

Using the generators in Σ_D and the relations in R_0, it is possible to eliminate all two-level operators of type X. As a first step, it must be shown that the two-level operators can be decomposed into circuits over Σ_D. This follows from the fact that Σ_D generates $W(E_8)$, and $\mathrm{S}(n) \hookrightarrow W(E_8)$.

$$X_{[0,1]} = X_0 \circ X_1 \circ CCX_{0,1} \circ X_1 \circ X_0 \qquad X_{[1,2]} = X_0 \circ CCX_{0,1} \circ CCX_{0,2} \circ CCX_{0,1} \circ X_0$$
$$X_{[2,3]} = X_0 \circ CCX_{0,1} \circ X_0 \qquad X_{[3,4]} = X_0 \circ X_2 \circ CCX_{0,1} \circ X_0 \circ CCX_{1,2} \circ CCX_{0,2} \circ CCX_{1,2} \circ X_0 \circ CCX_{0,1} \circ X_2 \circ X_0$$
$$X_{[4,5]} = X_1 \circ CCX_{0,1} \circ X_1 \qquad X_{[5,6]} = CCX_{0,1} \circ CCX_{0,2} \circ CCX_{0,1}$$

These relations can be validated with respect to $[\![\cdot]\!]_O$, and consequently introduced via a sequence of **Gen**(+) operations. These definitional relations can then be used to eliminate the two-level operators of type X, as outlined in Appendix B.2.

In this new presentation, then relations Relations (47), (56) and (57) are replaced by relations over Σ_D. Since R_0 is complete for $W(E_8)$, then these relations can be derived from R_0. Consequently, these relations can be eliminated with a sequence of **Rel**(−) transformations. This yields a new set of relations, denoted $\mathscr{R}_8^4$, which can be found in Figure 7. For simplicity of presentation, we used $X_{[0,1]}$ through to $X_{[6,7]}$ as abbreviations for the circuits given above. Furthermore, we take $CCZ = (-1)_{[7]}$ to be a generator with $(-1)_{[0]}$ an alias for $X_0 \cdot X_1 \cdot CCZ \cdot X_1 \cdot X_0$. Then define $\Sigma_1 = \Sigma_D \cup \{K_{[0,1,2,3]}, CCZ\}$ and $R_1 = R_0 \cup \mathscr{R}_8^4$, where $K_{[0,1,2,3]}$ corresponds to a negatively controlled K gate.

Theorem 5.2. $\mathrm{O}(8,\mathbb{D}) \cong \langle \Sigma_1 \mid R_1 \rangle$.

$$H_2 \cdot X_0 \approx X_0 \cdot H_2 \quad (102)$$
$$H_2 \cdot CX_{0,1} \approx CX_{0,1} \cdot H_2 \quad (103)$$
$$H_2 \cdot CCX_{1,2} \approx K_{0,1} \cdot K_{1,2} \cdot CCZ \cdot K_{1,2} \cdot K_{0,1} \cdot H_2 \quad (104)$$
$$H_2 \cdot CCZ \approx CCX_{0,1} \cdot H_2 \quad (105)$$
$$H_2 \cdot K_{1,2} \approx K_{1,2} \cdot H_2 \quad (106)$$
$$H_2 \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot H_2 \quad (107)$$
$$H_2 \cdot H_2 \approx \varepsilon \quad (108)$$

Figure 8: Additional relations for $\mathsf{TofH}(3)$.

5.7 A Minimal Generating Set for $\mathrm{O}(8,\mathbb{D})$

It turns out that CCZ and $K_{[0,1,2,3]}$ can be defined in terms of one-another, given that generators in Σ_D. The decompositions are as follows.

$$CCZ = K_{1,2} \circ CZ_{1,2} \circ X_0 \circ K_{[0,1,2,3]} \circ X_0 \circ CZ_{1,2} \circ K_{[0,1,2,3]} \circ X_{[5,6]} \qquad K_{[0,1,2,3]} = (K_{1,2} \circ CCZ)^3 \circ X_{[5,6]}$$

Given this observation, it seems natural to eliminate the $K_{[0,1,2,3]}$, given that it is not a common generator in quantum computation. However, the set $\Sigma_D \cup \{CCZ\}$ is minimal, whereas the set $\Sigma_D \cup \{K_{[0,1,2,3]}\}$ is not. In other words, choosing the generator $K_{[0,1,2,3]}$ enables a smaller generating set, whereas choosing the generator CCZ allows for more conventional circuit decompositions. For this reason, we choose to keep both $K_{[0,1,2,3]}$ and CCZ in our presentation. The minimality of these generating sets are proven in Appendix E, using the same techniques as in Section 4.5.

Theorem 5.3. *The following generating sets are minimal for* $\mathrm{O}(8,\mathbb{D})$.

1. $\Sigma_K = \{X_0, CX_{0,1}, CCX_{1,2}, K_{[0,1,2,3]}\}$.
2. $\Sigma_Z = \{X_0, CX_{0,1}, CCX_{1,2}, K_{1,2}, CCZ\}$

6 Extending to the 3-Qubit Toffoli-Hadamard Circuits

We now give a presentation of $\mathsf{TofH}(3)$, by leveraging the presentation of $\mathrm{O}(8,\mathbb{D})$ found in Section 5. The argument in this section closely follows [18, Section 5]. From [3], it is known that $\mathsf{TofH}(3)$ is obtained by adding the generator H_2 to $\mathrm{O}(8,\mathbb{D})$. Let $\Sigma_2 = \Sigma_1 \cup \{H_2\}$ and R_2 extend the set R_1 with all relations found in Figure 8. Using the relations in Figure 8, the generator H_2 can be moved from the left-hand side to the right-hand side of any word over Σ_1. Since H_2 is self-inverse, this is sufficient to decide equality in $\mathsf{TofH}(3)$.

The proof proceeds as follows. In Lemma 6.1, it is shown that H_2 commutes with every word in over Σ_1 using only the relations in R_2. This is used in Lemma 6.2, to show that every word over Σ_2 can be rewritten as a word over Σ_1, followed by at most one H_2 gate. Since $R_1 \subseteq R_2$ is a complete equational theory for $\mathrm{O}(8,\mathbb{D})$ with every element of $\mathsf{TofH}(3)$ of the form described in Lemma 6.2, it follows that R_2 is a complete equational theory for $\mathsf{TofH}(3)$ (see Theorem 6.3).

Lemma 6.1. *If* $w \in \Sigma_1^*$, *then there exists a* $w' \in \Sigma_1^*$ *such that* $H_2 \cdot w \sim_{R2} w' \cdot H_2$.

Proof. The proof follows by induction on $|w|$.

- **Base Case**. If $|w| = 0$, then $H_2 \cdot w = w \cdot H_2$. Then $H_2 \cdot w \sim_{R2} w \cdot H_2$ by the transitivity of $(\sim_{R2})$.
- **Inductive Case**. Assume that for some $k \in \mathbb{N}$, if $|w| = k$, then there exists a $w' \in \Sigma_1^*$ such that $H_2 \cdot w \sim_{R2} w' \cdot H_2$.

– **Inductive Step**. Assume that $|w| = k+1$. Then there exists a $u \in \Sigma_1^*$ and $x \in \Sigma_1$ such that $x \cdot u = w$. It follows by one of Relations (102), (103), (104), (105), (106) and (107), that there exits a $u' \in \Sigma_1^*$ such that $H_2 \cdot w \sim_{R2} u' \cdot H_2 \cdot u$. Since $|u| = k$, then by the inductive hypothesis, there exists a $w' \in \Sigma_1^*$ such that $H_2 \cdot u \sim_{R2} w' \cdot H_2$. Then $u' \cdot H_2 \cdot u \sim_{R2} u' \cdot w' \cdot H_2$. Then $H_2 \cdot u \sim_{R2} u' \cdot w' \cdot H_2$ by the transitivity of $(\sim_{R2})$, and the inductive case holds.

Then by the principle of induction, there exists a $w' \in \Sigma_1^*$ such that $H_2 \cdot w \sim_{R2} w' \cdot H_2$. □

Lemma 6.2. *If $w \in \Sigma_2^*$, then there exists some $w' \in \Sigma_1^*$ and $\ell \in \{0,1\}$ such that $w \sim_{R_2} w' \cdot H_2^\ell$.*

Proof. Let $f : \Sigma_2^* \to \mathbb{N}$ map each word $w \in \Sigma_2^*$ to the number of H_2 symbols in w. The proof follows by induction on $f(w)$.

– **Base Case**. If $f(w) = 0$, then $w \in \Sigma_1^*$. Then $w \sim_{R2} w \cdot H_2^0$ by the reflexivity of $(\sim_{R2})$

– **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $f(w) = k$, then there exists some $w' \in \Sigma_1^*$ and $\ell \in \{0,1\}$ such that $w \sim_{R2} w' \cdot H_2^\ell$.

– **Inductive Step**. Assume that $f(w) = k+1$. Then there exists $w_1 \in \Sigma_2^*$ and $w_2 \in \Sigma_1^*$ such that $w = w_1 \cdot H_2 \cdot w_2$ with $f(w_1) = f(w) - 1$. Then by Lemma 6.1, $w \sim_{R2} w_1 \cdot w_2' \cdot H_2$ for some $w_2' \in \Sigma_1^*$. Then $f(w_1 \cdot w_2') = f(w_1) = f(w) - 1$. By the inductive hypothesis, there exists some $w_3 \in \Sigma_1^*$ and $\ell \in \{0,1\}$ such that $w_1 \cdot w_2' \sim_{R2} w_3 \cdot H_2^\ell$. Then $w \sim_{R2} w_3 \cdot H_2^{\ell+1}$. If $\ell = 0$, then $w \sim_{R2} w_3 \cdot H_2$ and we are done. Otherwise, if $\ell = 1$, then $w \sim_{R2} w_3$ by Relation (108). In either case, there exists an $\ell' \in \{0,1\}$ such that $w \sim_{R2} w_3 \cdot H_2^{\ell'}$ and the inductive step holds.

Then by the principle of induction, there exists some $w' \in \Sigma_1^*$ and $\ell \in \{0,1\}$ such that $w \sim_{R_2} w' \cdot H_2^\ell$. □

Theorem 6.3. *For all $w_1, w_2 \in \Sigma_2^*$, $[\![w_1]\!]_H^* = [\![w_2]\!]_H^*$ if and only if $w_1 \sim_{R2} w_2$.*

Proof. It follows by matrix multiplication that the relations in Figure 8 are sound. It remains to be shown that the relations in Figure 8 are complete. Let $w_1 \in \Sigma_2^*$ and $w_2 \in \Sigma_2^*$ such that $[\![w_1]\!]_H^* = [\![w_2]\!]_H^*$. By Lemma 6.2, there exists $\ell_1, \ell_2 \in \{0,1\}$ and $w_1', w_2' \in \Sigma_1^*$ such that $w_1 \sim_{R2} w_1' \cdot H_2^{\ell_1}$ and $w_2 \sim_{R2} w_2' \cdot H_2^{\ell_2}$. Since $[\![w_1]\!]_H^* \in \mathrm{O}(8,\mathbb{D}) \cong \langle \Sigma_1, R_1 \rangle$ with $R_1 \subseteq R_2$, then $w_1' \cdot H_2^{\ell_1} \sim_{R2} w_2' \cdot H_2^{\ell_1}$. Assume for the intent of contradiction that $\ell_1 \neq \ell_2$. Then $[\![w_1']\!]_H^* = [\![w_2']\!]_H^* \circ [\![H_2]\!]_H$. Then $[\![w_2']\!]_H^* \circ [\![H_2]\!]_H^* \in \mathrm{O}(8,\mathbb{D})$. However, $[\![w_2']\!]_H^* \circ [\![H_2]\!]_H$ has a denominator of the form $1/(2^k\sqrt{2})$, and therefore $[\![w_2']\!]_H^* \circ [\![H_2]\!]_H \notin \mathrm{O}(8,\mathbb{D})$. By contradiction, $\ell_1 = \ell_2$. Since $\ell_1 = \ell_2$, then $w_1 \sim_{R2} w_2$ by the transitivity and symmetry of $(\sim_{R2})$. Since w_1 and w_2 were arbitrary, then the relations in Figure 8 are complete. □

7 Conclusion

We used the geometry of $W(E_8)$ to obtain a circuit presentation for the 3-qubit Toffoli-K circuits, and then leveraged [18] to obtain a finite presentation of 3-qubit Toffoli-Hadamard circuits. Our presentation contains 65 relations, compared to the 2113 relations of [18]. There are several directions for future work. We hope to simplify our presentation by further reducing the number of relations. In addition, we plan to explore the structural properties of the group of 3-qubit Toffoli-Hadamard circuits and of its subgroups. In particular, it is known that the group $\mathrm{O}(8,\mathbb{D})$ is generated by reflections, but it is not known whether this group can be presented as an (infinite) Coexeter group. From an applied perspective, we also hope to explore applications of these presentations to circuit optimization and equivalence checking.

References

[1] Dorit Aharonov (2003): *A simple proof that Toffoli and Hadamard are quantum universal*. arXiv:quant-ph/0301040.

[2] Matthew Amy, Jianxin Chen & Neil J. Ross (2018): *A Finite Presentation of CNOT-Dihedral Operators*. *EPTCS* 266, pp. 84–97, doi:10.4204/eptcs.266.5.

[3] Matthew Amy, Andrew Glaudell & Neil Ross (2020): *Number-Theoretic Characterizations of Some Restricted Clifford+T Circuits*. *Quantum* 4, p. 252, doi:10.22331/q-2020-04-06-252.

[4] Matthew Amy, Andrew N. Glaudell, Sarah Meng Li & Neil J. Ross (2023): *Improved Synthesis of Toffoli-Hadamard Circuits*. In: *Reversible Computation*, Springer-Verlag, pp. 169–209, doi:10.1007/978-3-031-38100-3_12.

[5] Matthew Amy, Neil J. Ross & Scott Wesley (2024): *Supplement: A Sound and Complete Equational Theory for 3-Qubit Toffoli-Hadamard Circuits*. Available as an ancillary file from the arXiv page of this paper.

[6] Franz Baader & Tobias Nipkow (1998): *Term Rewriting and All That*. Cambridge University Press, doi:10.1017/CBO9781139172752.

[7] Xiaoning Bian & Peter Selinger (2023): *Generators and relations for 2-qubit Clifford+T operators*. *EPTCS* 394, pp. 13–28, doi:10.4204/eptcs.394.2.

[8] Xiaoning Bian & Peter Selinger (2023): *Generators and relations for 3-qubit Clifford+CS operators*. *EPTCS* 384, pp. 114–126, doi:10.4204/eptcs.384.7.

[9] Ronald V. Book & Friedrich Otto (1993): *String-Rewriting Systems*. Springer, doi:10.1007/978-1-4613-9771-7.

[10] Maria Luisa Dalla Chiara, Antonio Ledda, Giuseppe Sergioli & Roberto Giuntini (2013): *The Toffoli-Hadamard gate system: an algebraic approach*. *Journal of Philosophical Logic* 42, pp. 467–481, doi:10.1007/s10992-013-9271-9.

[11] Alexandre Clément, Nicolas Heurtel, Shane Mansfield, Simon Perdrix & Benoît Valiron (2023): *A Complete Equational Theory for Quantum Circuits*. In: *LiCS*, IEEE, pp. 1–13, doi:10.1109/LICS56636.2023.10175801.

[12] J. H. Conway & N. J. A. Sloane (1987): *Sphere-Packings, Lattices, and Groups*. Springer-Verlag, doi:10.1007/978-1-4757-6568-7.

[13] Jörg Endrullis, Herman Geuvers, Jakob Grue Simonsen & Hans Zantema (2011): *Levels of undecidability in rewriting*. *Information and Computation* 209(2), pp. 227–245, doi:10.1016/j.ic.2010.09.003.

[14] Adam P. Goucher (2020): *Minimalistic Quantum Computation*. `https://cp4space.hatsya.com/2020/05/10/minimalistic-quantum-computation/`. Accessed: 2023-11-26.

[15] Simon Henry & Samuel Mimram (2022): *Tietze Equivalences as Weak Equivalences*. *Applied Categorical Structures* 30(3), pp. 453–483, doi:10.1007/s10485-021-09662-w.

[16] James E. Humphreys (1990): *Reflection Groups and Coxeter Groups*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, doi:10.1017/CBO9780511623646.

[17] D. L. Johnson (1990): *Presentations of Groups*. Cambridge University Press, doi:10.1017/CBO9781139168410.

[18] Sarah Meng Li, Neil J. Ross & Peter Selinger (2021): *Generators and Relations for the Group $On(Z[1/2])$*. *EPTCS* 343, pp. 210–264, doi:10.4204/eptcs.343.11.

[19] Justin Makary, Neil J. Ross & Peter Selinger (2021): *Generators and Relations for Real Stabilizer Operators*. *EPTCS* 343, p. 14–36, doi:10.4204/eptcs.343.2.

[20] Leonardo de Moura & Nikolaj Bjørner (2008): *Z3: An Efficient SMT Solver*. In: *TACAS*, Springer-Verlag, pp. 337–340, doi:10.5555/1792734.1792766.

[21] Michel Planat (2011): *Clifford group dipoles and the enactment of Weyl/Coxeter group $W(E8)$ by entangling gates*. *Gen. Math. Notes* 2(1), pp. 96–113, doi:10.22331/q-2020-04-06-252.

[22] Peter Selinger (2015): *Generators and Relations for n-Qubit Clifford Operators*. *LMCS* 11(2:10), pp. 1–17, doi:10.2168/LMCS-11(2:10)2015.

[23] Yaoyun Shi (2003): *Both Toffoli and controlled-NOT need little help to do universal quantum computing*. *Quantum Info. Comput.* 3(1), pp. 84–92, doi:10.5555/2011508.2011515.

[24] Maryna S. Viazovska (2017): *The sphere packing problem in dimension 8*. *Annals of Mathematics* 185(3), pp. 991–1015, doi:10.4007/annals.2017.185.3.7.

[25] Renaud Vilmart (2019): *A ZX-calculus with triangles for Toffoli-Hadamard, Clifford+T, and beyond*. *EPTCS* 287, pp. 313–344, doi:10.4204/eptcs.287.18.

[26] Renaud Vilmart (2023): *Completeness of sum-over-paths for Toffoli-Hadamard and the dyadic fragments of quantum computation*. In: *CSL*, *LIPIcs* 252, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 36:1–36:17, doi:10.4230/LIPIcs.CSL.2023.36.

A Semantic Tietze Transformations

This section recalls what it means for a function to induce a monoid homomorphism. This is then used to prove the soundness and completeness of Tietze transformations with respect to semantic interpretations. The majority of this section is dedicated to proving that all induced homomorphisms are isomorphisms, and in the case of **Gen(+)**, the extension is unique. The uniqueness of this extension is necessary to prove that each generator in Σ_D has the intended matrix semantics.

A.1 Induced Monoid Homomorphisms

Let Σ be an alphabet and M a monoid. For each function $f : \Sigma \to M$, define the function $f^* : \Sigma^* \to M$ such that $f^*(x_1 \cdot x_2 \cdots x_n) = 1_M \circ f(x_1) \circ f(x_2) \circ \cdots \circ f(x_n)$ for all $x_1 \cdot x_2 \cdots x_n \in \Sigma^*$. It can then be shown that f^* is the unique monoid homomorphism such that $f^*(x) = f(x)$ for all $x \in \Sigma$ [17]. Given a set of relations $R \subseteq \Sigma^* \times \Sigma^*$, it can then be asked whether f induces a monoid homomorphism between $G = \langle \Sigma \mid R \rangle$ and M. This question is answered by the following theorem.

Theorem A.1 ([9])**.** *Let M and $G = \langle \Sigma \mid R \rangle$ be monoids with $f : \Sigma \to M$. There there exists a unique monoid homomorphism $\varphi : G \to M$ such that $f^* = \varphi \circ \pi_G$ if and only if $f^*(q) = f^*(r)$ for all $q \approx_R r$. In this case,* $\mathrm{Im}(\varphi) = \langle \varphi(\pi_G(\Sigma)) \rangle$.

Theorem A.1 characterizes when f induces a monoid homomorphism, and how to construct this induced homomorphism φ. It can then be asked how the construction of φ interacts with the introduction or elimination of generators. As outlined by the following theorems, the elimination of generators corresponds to certain restrictions of φ, whereas the introduction of generators corresponds to certain unique extensions of φ.

Lemma A.2. *Let Σ be an alphabet, $x \in \Sigma$, $\Sigma' = \Sigma \setminus \{x\}$, and $D = \{x \approx w\}$ for some $w \in (\Sigma')^*$. If $q \in \Sigma^*$, then there exists a $q' \in (\Sigma')^*$ such that $q \sim_D q'$.*

Proof. Let $f : \Sigma^* \to \mathbb{N}$ count the x symbols in a word. Then the proof follows by induction on $f(q)$.

- **Base Case**. If $f(q) = 0$, then $q \in (\Sigma')^*$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $f(q) = k$, then there exists some $q' \in (\Sigma')^*$ such that $q \sim_D q'$.
- **Inductive Step**. Assume $f(q) = k+1$. Since $f(q) > 1$, then there exists some $u, v \in \Sigma^*$ such that $q = u \cdot x \cdot v$. Since $k+1 = f(q) = f(u \cdot x \cdot v) = f(u) + f(x) + f(v) = f(u) + f(v) + 1$, then $k = f(u) + f(v)$. Since $x \approx_D w$, then $q \sim_D u \cdot w \cdot v$ with $f(u \cdot w \cdot v) = f(u) + f(v) = k$. Then by the inductive hypothesis, there exists some $q' \in (\Sigma')^*$ such that $u \cdot w \cdot v \sim_D q'$. Then $q \sim q'$.

By the principle of induction, there exists some $q' \in (\Sigma')^*$ such that $q \sim_D q'$. □

Theorem A.3. *Let M and $G = \langle \Sigma \mid R \rangle$ be monoids with $f : \Sigma \to M$, and $H = \langle \Sigma \cup \{x\} \mid R \cup \{x \approx w\} \rangle$ be a monoid for some $x \notin \Sigma$ and $w \in \Sigma^*$. Define $g : \Sigma \cup \{x\} \to M$ such that $g|_\Sigma = f$ and $g : x \mapsto f^*(w)$. If f induces a monoid homomorphism from G to M, then g is the unique extension of f to induce a monoid homomorphism from H to M. Furthermore, if f induces an injection (resp. surjection) from G to M, then g induces an injection (resp. surjection) from H to M.*

Proof. Let $\Pi = \Sigma \cup \{x\}$ and $Q = R \cup \{x \approx w\}$. Assume that f induces a homomorphism from G to M.

– (**Induced Hom**). Let $q \approx_Q r$. Then either $q \approx_R r$ or $(q,r) = (x,w)$. First, assume that $q \approx_R r$. Then $f^*(q) = f^*(r)$ by Theorem A.1. Then $g^*(q) = (g|_\Sigma)^*(q) = f^*(q) = f^*(r) = (g|_\Sigma)^*(r) = g^*(r)$. Next, assume that $(q,r) = (x,w)$. Then $g^*(x) = g(x) = f^*(w) = (g|_\Sigma)^*(w) = g^*(w)$. In either case $g^*(q) = g^*(r)$. Since $q \approx_R r$ was arbitrary, then $g^*(q) = g^*(r)$ for all $q \approx_R r$. Then g induce a monoid homomorphism from H to M by Theorem A.1.

– (**Uniqueness**). Assume that $k : \Sigma \cup \{x\} \to M$ is an extension of f which induces a monoid homomorphism. Since k induces a monoid homomorphism and $x \approx_Q w$, then $k^*(x) = k^*(w)$ by Theorem A.1. Then $k(x) = k^*(x) = k^*(w) = (k|_\Sigma)^*(w) = f^*(w) = g(x)$ by construction of g. Moreover, since $k|_\Sigma = f = g|_\Sigma$, then $k = g$. Since k was arbitrary, then g is unique.

– (**Injectivity**). Let f induce φ and g induce ρ. Assume that φ is injective. Let $q,r \in \Pi^*$ such that $\rho(\pi_H(q)) = \rho(\pi_H(r))$. Since $x \approx_Q w$, then by Lemma A.2 there exists $q',r' \in \Sigma^*$ such that $q \sim_Q q'$ and $r \sim_Q r'$. Then $\pi_H(q) = \pi_H(q')$ and $\pi_H(r) = \pi_H(r')$. Then $\rho(\pi_H(q')) = \rho(\pi_H(r'))$. Then $g^*(q') = g^*(r')$. Since $q',r' \in \Sigma^*$, then $f^*(q') = f^*(r')$. Then $\varphi(\pi_G(q')) = \varphi(\pi_G(r'))$. Since φ is injective, then $\pi_G(q') = \pi_G(r')$. Then $q' \sim_R r'$. Since $R \subseteq Q$, then $q' \sim_Q r'$. Then $\pi_H(q') = \pi_H(r')$. Then $\pi_H(q) = \pi_H(r)$. Since q and r were arbitrary, then ρ is injective.

– (**Surjectivity**). Let f induce φ and g induce ρ. Assume that φ is surjective. Since $\varphi(\pi_G(y)) = f^*(y) = g^*(y) = \rho(\pi_H(y))$ for each $y \in \Sigma$, then $\varphi(\pi_G(\Sigma)) \subseteq \mathrm{Im}(\rho)$. Since φ is surjective, then $H = \mathrm{Im}(\varphi) = \langle \varphi(\pi_G(\Sigma)) \rangle \leq \mathrm{Im}(\rho) \leq H$ and ρ is surjective.

Therefore, g is the unique extension of f to induce a monoid homomorphism from H to M, with g inducing an injection (resp. surjection) whenever f induces an injection (resp. surjection). □

Theorem A.4. *Let M and $G = \langle \Sigma \mid R \rangle$ be monoids with $f : \Sigma \to M$ and $H = \langle \Pi \mid Q \rangle$ where $\Pi = \Sigma \setminus \{x\}$ for some $x \in \Sigma$ and $Q = R \setminus \{x \approx w\}$ for some $x \approx_R w$. If $Q \subseteq \Pi^* \times \Pi^*$ and f induces a monoid homomorphism from G to M, then $f|_\Pi$ induces a monoid homomorphism from H to M. Furthermore, if f induces an injection (resp. surjection) from G to M, then $f|_\Pi$ induces an injection (resp. surjection) from H to M.*

Proof. Assume that f induces a homomorphism from G to M.

– (**Induced Hom**). Since f induces a monoid homomorphism, then $f^*(q) = f^*(r)$ for all $q \approx_R r$ by Theorem A.1. Since $Q \subseteq R$ and $Q \subseteq \Pi^* \times \Pi^*$, then $f|_\Pi^*(q) = f^*(q) = f^*(r) = f|_\Pi^*(r)$ for all $q \approx_Q r$. Then $f|_\Pi$ induces a monoid homomorphism from H to M by Theorem A.1.

– (**Injectivity**). Let f induce φ and $f|_\Pi$ induce ρ. Assume that φ is injective. Let $q,r \in \Pi^*$ such that $\rho(\pi_H(q)) = \rho(\pi_H(r))$. Then $f|_\Pi^*(q) = f|_\Pi^*(r)$. Then $f^*(q) = f^*(r)$. Then $\varphi(\pi_G(q)) = \varphi(\pi_G(r))$. Since φ is an injective, then $\pi_G(q) = \pi_G(r)$. Then $q \sim_R r$. Since $Q = R \setminus \{x \approx w\}$, $Q \subseteq \Pi^* \times \Pi^*$, and $q,r \in \Pi^*$, then $q \sim_Q r$. Then $\pi_H(q) = \pi_H(r)$. Since q and r were arbitrary, then ρ is injective.

– (**Surjectivity**). Let f induce φ and $f|_\Pi$ induce ρ. Assume that φ is surjective. Since $\varphi(\pi_G(y)) = f^*(y) = f|_\Pi^*(y) = \rho(\pi(y)) \in \mathrm{Im}(\rho)$ for each $y \in \Pi$, then $\varphi(\pi_G(\Pi)) \subseteq \mathrm{Im}(\rho)$. Since $x \approx_R w$, then $\pi_G(x) = \pi_G(w)$, and consequently $\varphi(\pi_G(x)) = \varphi(\pi_G(w)) = f^*(w) = f|_\Pi^*(w) = \rho(\pi_H(w)) \in \mathrm{Im}(\rho)$. Then $\varphi(\pi_G(\Sigma)) \subseteq \mathrm{Im}(\rho)$. Since φ is surjective, then $H = \mathrm{Im}(\varphi)\langle \varphi(\pi_G(\Sigma)) \leq \mathrm{Im}(\rho) \leq H$ and ρ is surjective.

Therefore, $f|_\Pi$ induces a monoid homomorphism from H to M with $f|_\Pi$ inducing an injection (resp. surjection) whenever f induces an injection (resp. surjection). □

A.2 Semantic Interpretations and Relations

Let $G = \langle \Sigma \mid R \rangle$ be a monoid presentation with an interpretation $[\![\cdot]\!]_G : \Sigma \mapsto H$. The **Rel(+)** transformation states that if $r \in \Sigma^*$ and $q \in \Sigma^*$ with $r \sim_R q$, then $\langle \Sigma \mid R \rangle \cong \langle \Sigma \mid R \cup \{r\} \rangle$. In practice, deriving q from r can be challenging, and on a theoretical level, this is known to be undecidable [13]. However, it is rarely the case that one would try to prove $r \sim_R q$ without some intuition that $\pi_G(r) = \pi_G(q)$. In the case of this paper, this intuition comes from knowledge about operators in $O(8, \mathbb{D})$. For example, if $M \circ N = A \circ B$, then for any complete set of relations R, it must be the case that $M \cdot N \sim_R A \cdot B$. More generally, if $[\![r]\!]^*_\Sigma = [\![q]\!]^*_\Sigma$ with $[\![\cdot]\!]^*_\Sigma$ inducing an injection, then $r \sim_R q$. This claim is established by the following theorem, and used freely throughout the paper to simplify derivations.

Definition A.5 (Valid Semantic Interpretation)**.** A semantic interpretation $[\![\cdot]\!]_\Sigma : G \to H$ for a presentation $G = \langle \Sigma \mid R \rangle$ is valid if $[\![q]\!]^*_\Sigma = [\![r]\!]^*_\Sigma$ for all $r \approx_R q$.

Theorem A.6. *Let $G = \langle \Sigma \mid R \rangle$ be a presentation with a valid semantic interpretation $[\![\cdot]\!]_G : \Sigma \mapsto H$. If $[\![\cdot]\!]_G$ is injective and $[\![q]\!]^*_\Sigma = [\![r]\!]^*_\Sigma$, then $q \sim_R r$.*

Proof. Since $[\![\cdot]\!]_\Sigma$ is valid, then by Theorem A.1, $[\![\cdot]\!]_\Sigma$ induces a monoid homomorphism $\varphi : G \to H$ such that $[\![\cdot]\!]^*_\Sigma = \varphi \circ \pi_G$. Assume that $[\![q]\!]^*_\Sigma = [\![r]\!]^*_\Sigma$. Then $\varphi(\pi_G(q)) = [\![q]\!]^*_\Sigma = [\![r]\!]^*_\Sigma = \varphi(\pi_G(r))$. Since φ is injective, then $\pi_G(q) = \pi_G(r)$. Then $q \sim_R r$. □

A.3 Semantics and Generator Introduction

In the previous section, it was assumed that $[\![\cdot]\!]_\Sigma$ inducted an injection. This is a reasonable assumption. For example, if $G = \langle \Sigma \mid R \rangle$ is a presentation for H, then there exists an isomorphism $G \cong H$ from which $[\![\cdot]\!]_\Sigma$ can be extracted.

This can become problematic when trying to translate a known presentation to a desired generator set. Assume that $\langle \Pi \mid Q \rangle$ is a known presentation with a semantic interpretation $[\![\cdot]\!]_\Pi : \Pi \to H$, from with a presentation $\langle \Sigma \mid R \rangle$ is derivable via a sequence of Tietze transformations. One would hope that after each Tietze transformation, the semantic interpretation continues to induce an injection, so that Theorem A.6 continues to hold. Furthermore, one would hope that after all of the Tietze transformations, $[\![\cdot]\!]_\Sigma$ is a valid semantic interpretation.

It will be shown that under reasonable assumptions, all Tietze transformations satisfy these assumptions. The first concern is answered by Theorem A.3 and Theorem A.4, which state that after each **Gen(+)** or **Gen(−)** transformation, $[\![\cdot]\!]_\Pi$ continues to induce a monoid homomorphism (resp. injection, surjection, isomorphism). The only time injectivity might fail is after a new generator is added. In **Gen(+)**, the symbol $x \in \Sigma$ becomes an alias for $w \in \Pi^*$. Then it suffices to check that $[\![x]\!]_\Sigma = [\![w]\!]^*_\Pi$.

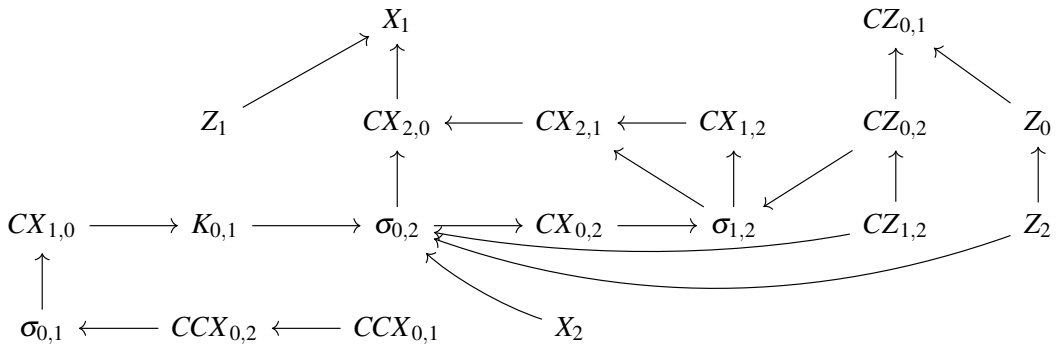

Figure 9: The derived generator graph for the defining relations in R_0.

B From Derivations to Tietze Transformations

This section describes higher-level structures to reason about Tietze transformations. It is proven that each structure corresponds to a valid sequence of Tietze transformations, and is therefore sound for the isomorphism of finite monoid presentations. The structures described in this section are used throughout Section 4 and Section 5.

B.1 Digraphs and Termination

This section reviews the basics of directed graphs. A directed graph is a tuple (V,E) such that $V \subseteq E \times E$. A vertex v is a *child* of u if $(u,v) \in E$. A vertex v is a *parent* of u if $(v,u) \in E$. A *path* in (V,E) is some sequence $(u_0,u_1,\ldots,u_n)$ over V such that $(u_{k-1},u_k) \in E$ for all $k \in \{1,2,\ldots,n\}$. A directed graph (V,E) is *acyclic* if $u_0 \neq u_n$ for all paths $(u_0,u_1,\ldots,u_n)$ in (V,E).

Lemma B.1 ([6, Sect. 2.2])**.** *If a digraph (V,E) is finite and acyclic, then for every vertex $v \in V$, there exists some path of length n ending (resp. starting) at v such that every path ending (resp. starting) at v has length at most n. In particular, there exists a vertex $v \in V$ such that v has no children (resp. parents) in (V,E).*

Note that the original statement of Lemma B.1 concerned the termination of abstract rewriting systems (which can be modelled using paths through digraphs). In the proofs that follow, this intuition is useful to keep in mind. Indeed, Lemma B.1 is used to argue that certain rewriting procedures terminate, though the rewriting systems are never stated explicitly for simplicity.

B.2 Derived Generators and Tietze Transformations

In Section 4.1, it was claimed that 19 of the generators in Σ_D could be introduced freely, because their defining relations formed an acyclic dependency graph. Later, in Appendix E, the same argument was used to remove these 19 generators from the generating set for $W(E_8)$. Similar arguments appear throughout Section 5. The goal of this section is to establish both directions rigorously. First, formal definitions are given for defining relations and derived generator graphs (i.e., the graph of dependencies between the relations). Then, Theorem B.6 and Theorem B.11 are established to justify the claims of Section 4.1 and Section 4.5 respectively. Finally, it is shown that these theorems apply to the generators in Section 4.

Definition B.2 (Defining Relations)**.** Let Σ be an alphabet. A Σ-*defining relation* for $x \in \Sigma$ is a relation of the form $x \approx w$ where $w \in \Sigma^*$. A *family of* Σ-*defining relations for* $\Pi \subseteq \Sigma$ is a set $\{r_x \mid x \in \Pi\}$ such that for each $x \in \Pi$, r_x is a Σ-defining relation for x.

Definition B.3 (Derived Generator Graph)**.** Let D be a family of Σ-defining relations for $\Pi \subseteq \Sigma$. The *derived generator graph for* D is the digraph $\Gamma_D(D) := (\Pi, E)$ such that $(x,y) \in E$ if and only if there exists $w, w' \in \Sigma^*$ such that $x \approx w \cdot y \cdot w'$ is a relation in D.

The proofs of Theorem B.6 and Theorem B.11 both rely heavily on derived generator graphs. Intuitively, a generator can be introduced (resp. eliminated) if it has no dependencies (resp. dependants) in the derived generator graph. It is always possible to find such a generator, provided the graph is acyclic. However, once a generator has been removed from the defining relations, it is important that the new derived generator graph is also acyclic. In fact, the new derived generator graph is always a subgraph of the previous derived generator graph, as outlined in Lemma B.4.

Lemma B.4. *If D is a family of Σ-defining relations and $D' \subseteq D$, then $\Gamma_D(D')$ is a subgraph of $\Gamma_D(D)$.*

Proof. If D is a family of Σ-defining relations for Π, then there exists $\Pi' \subseteq \Pi$ such that D' is a family of Σ-defining relations for Π'. Then $\Gamma_D(D) = (\Pi, E)$ and $\Gamma_D(D') = (\Pi', E')$ for some $E \subseteq \Pi \times \Pi$ and $E' \subseteq \Pi' \times \Pi$. Let $(x,y) \in E'$, Then there exists some $w, w' \in \Sigma^*$ such that $x \approx_{D'} w \cdot y \cdot w'$. Since $D' \subseteq D$, then $x \approx_D w \cdot y \cdot y'$. Then $(x,y) \in E$. Since (x,y) was arbitrary, then $E' \subseteq E$. Since $\Pi' \subseteq \Pi$ and $E' \subseteq E$, then $\Gamma_D(D')$ is a subgraph of $\Gamma_D(D)$. □

B.2.1 Introduction of Derived Generators

The goal of this section is to prove Theorem B.6. First, Lemma B.5 is introduced to prove that every finite set of defining relations with an acyclic graph must contain at least one defining relation $x \approx w$ whose right-hand side consists only of primitive generators. Since w is consists only of primitive generators, then it may be introduced by a **Gen**$(+)$ transformation. The proof then follows by induction on the number of defining relations, as outlined below.

Lemma B.5. *If D is a family of Σ-defining relations for Π with Π finite and $\Gamma_D(D)$ acyclic, then there exists a relation $x \approx w$ in D such that $w \in (\Sigma \setminus \Pi)^*$.*

Proof. Since Π is finite and $\Gamma_D(D)$ is acyclic, the by Lemma B.1, there exists some vertex $x \in \Pi$ such that x has no children in $\Gamma_D(D)$. Let $x \approx_D w$ be the Σ-defining relation for x in D with $n = |w|$. Let $k \in [n]$. Assume for the intent of contradiction that $w_k \in \Pi$. Then there exists some $w', w'' \in \Sigma^*$ such that $w = w' \cdot w_k \cdot w''$. Then w_k is a child of x by definition. However, x has no children by assumption. Then $w_k \notin \Pi$ by contradiction. Since k was arbitrary, then $w \in (\Sigma \setminus \Pi)^*$. □

Theorem B.6. *Let $\Sigma \subseteq \widehat{\Sigma}$ be an alphabet with $\Pi = \widehat{\Sigma} \setminus \Sigma$ finite and $R \subseteq \Sigma^* \times \Sigma^*$. If D is a family of $\widehat{\Sigma}$-defining relations for Π with $\Gamma_D(D)$ acyclic, then there exists a length $|\Pi|$ sequence of **Gen**$(+)$ transformations between $\langle \Sigma \mid R \rangle$ and $\langle \widehat{\Sigma} \mid R \cup D \rangle$.*

Proof. The proof follows by induction on $|\Pi|$.

- **Base Case**. Assume that $|\Pi| = 0$. Then $|D| = 0$ and $\langle \Sigma \mid R \rangle = \langle \Sigma' \mid R \cup D \rangle$. Then there exists a length 0 sequence of **Gen**$(+)$ transformations between $\langle \Sigma \mid R \rangle$ and $\langle \widehat{\Sigma} \mid R \cup D \rangle$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|\Pi| = k$ and $\Pi_D(D)$ is acyclic, then there exists a length k sequence of **Gen**$(+)$ transformations between $\langle \Sigma \mid R \rangle$ and $\langle \widehat{\Sigma} \mid R \cup D \rangle$.

– **Inductive Step**. Assume that $|\Pi| = k+1$ and $\Gamma_D(D)$ is acylic. Since Π is finite and $\Gamma_D(D)$ is acyclic, then by Lemma B.5 there exists a $\widehat{\Sigma}$-defining relation $x \approx w$ in D. Let $\Lambda = \Sigma \cup \{x\}$ and $Q = R \cup \{x \approx w\}$. Then $D \setminus \{r_x\}$ is a family of $\widehat{\Sigma}$-defining relations for $\Pi \setminus \{x\}$ and $Q \subseteq \Lambda^* \times \Lambda^*$. Since $\Gamma_D(D \setminus \{r_x\})$ is a subgraph of $\Gamma_D(D)$ by Lemma B.4 with $\Gamma_D(D)$ acyclic, then $\Gamma_D(D \setminus \{r_x\})$ is also acyclic. Since $|\Pi \setminus \{x\}| = |\Pi| - 1 = k$ and $\Gamma_D(D \setminus \{r_x\})$ is acyclic, then by the inductive hypothesis, there exists a length k sequence of **Gen**(+) transformations from $\langle \Lambda \mid Q \rangle$ to $\langle \widehat{\Sigma} \mid R \cup D \rangle$. Furthermore, $\langle \Sigma \mid R \rangle \cong \langle \Lambda \mid Q \rangle$ by **Gen**(+). Then there exists a length $k+1$ sequence of **Gen**(+) transformations between $\langle \Sigma \mid D \rangle$ and $\langle \widehat{\Sigma} \mid R \cup D \rangle$.

By the principle of induction, there exists a length $|\Pi|$ sequence of **Gen**(+) transformations between $\langle \Sigma \mid D \rangle$ and $\langle \Sigma' \mid R \cup D \rangle$. □

B.2.2 Tietze Transformations to Exchange Relations

The goal of this section is to introduce a technique necessary to prove that derived generators can be eliminated via finite sequences of Tietze transformations. Using the Tietze transformations discussed so far, it is possible to remove and introduce redundant relation. In practice, one often wishes to replace a relation $q \approx w$ with a relation $q' \approx w'$, where neither $q \approx w$ nor $q' \approx w'$ is redundant without the other relation. The following lemma gives a sufficient condition for when $q \approx w$ can be exchanged with $q' \approx w'$, and provides an upper bound on the number of Tietze transformations required to carry out the exchange.

Lemma B.7. *Let $G = \langle \Sigma \mid R \rangle$ be a presentation with $q \approx w$ in R and $Q = R \setminus \{q \approx w\}$. If there exists $q' \in \Sigma^*$ and $w' \in \Sigma^*$ such that $q \sim_Q q'$ and $w \sim_Q w'$, then there exists a finite sequence of Tietze transformations between G and $\langle \Sigma \mid Q \cup \{q' \approx w'\} \rangle$ of length between 1 and 2.*

Proof. Let $S = R \cup \{q' \approx w'\}$ and $T = Q \cup \{q' \approx w'\}$. If $q' \approx_R w'$, then $\langle \Sigma \mid R \rangle = \langle \Sigma \mid S \rangle$. Assume instead that $q \not\approx_R w$. Since $Q \subseteq R$, then $q \sim_R q'$ and $w \sim_R w'$. Then $q' \sim_R q$ by the symmetry of $(\sim_R)$. Since $q' \sim_R q$, $q \sim_R w$, and $w \sim_R w'$, then $q' \sim_R w'$ by transitivity. Then $\langle \Sigma \mid R \rangle \cong \langle \Sigma \mid S \rangle$ by a single **Rel**(+) transformation. In either case, there exists a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid S \rangle$ of length at most one. Since $Q \subseteq T$, then $q \sim_T q'$ and $w \sim_T w'$. Then $w' \sim_T w$ by the symmetry of $(\sim_T)$. Since $q \sim_T q'$, $q' \sim_T w'$, and $w' \sim_T w$, then $q \sim_T w$ by the transitivity of $(\sim_S)$. Since $S \neq T$, then $\langle \Sigma \mid S \rangle \cong \langle \Sigma \mid T \rangle$ by a single **Rel**(+) transformation. In conclusion, there exists a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid S \rangle$ of length at between 1 and 2. □

B.2.3 Elimination of Derived Generators

The goal of this section is to prove Theorem B.11. The first step in this proof is to show that the set of relations can be replaced by $Q \cup D$ where Q is a set of relations over the primitive generators and D is the set of defining relations. This follows by induction in Lemma B.9, where Lemma B.8 is used to find new relations of the form, and then Lemma B.7 is used to exchange the relations. The second step in this proof is to show that the derived generators can be eliminated through a finite sequence of Tietze transformations. This follows by induction in Theorem B.11, where Lemma B.10 is used to find a derived generator upon which non other derived generator depends. Such a generator is necessarily redundant, and may be eliminated via a **Rel**(−) transformation.

Lemma B.8. *Let $\Sigma \setminus \widehat{\Sigma}$ be an alphabet with $\Pi = \widehat{\Sigma} \setminus \Sigma$. If D is a family of $\widehat{\Sigma}$-defining relations for Π with $\Gamma_D(D)$ acyclic, then for each $w \in \Sigma^*$, there exists some $w' \in (\Sigma')^*$ such that $w \sim_D w'$.*

Proof. The proof follows by induction on the size of Π.

- **Base Case**. Assume that $|\Pi| = 0$ and $w \in \widehat{\Sigma}^*$. Then $\Pi = \varnothing$ and $w \in \Sigma^*$. Since $w \sim_D w$ by the symmetry of $(\sim_D)$, then there exists some $w' \in \Sigma^*$ such that $w \sim_D w'$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|\Pi| = k$, $\Gamma_D(D)$ is acyclic, and $w \in \widehat{\Sigma}^*$, then there exists some $w' \in \Sigma^*$ such that $w \sim_D w'$.
- **Inductive Step**. Assume that $|\Pi| = k+1$ and $\Gamma_D(D)$ is acyclic. Since Π is finite and $\Gamma_D(D)$ is acyclic, then by Lemma B.10, there exists a $\widehat{\Sigma}$-defining relation $x \approx q$ in D such that $D \setminus \{r_x\}$ is a family of $(\widehat{\Sigma} \setminus \{x\})$-defining relations for $\Pi \setminus \{x\}$. Then by Lemma A.2, there exists some $w' \in (\widehat{\Sigma} \setminus \{x\})^*$ such that $w \sim_D w'$. Since $\Gamma_D(D \setminus \{r_x\})$ is a subgraph of $\Gamma_D(D)$ by Lemma B.4 with $\Gamma_D(D)$ acyclic, then $\Gamma_D(D \setminus \{r_x\})$ is also acyclic. Then $D \setminus \{r_x\}$ is a family of $(\widehat{\Sigma} \setminus \{x\})$-defining relations for $\Pi \setminus \{x\}$ with $\Gamma_D(D)$ acyclic and $w' \in (\widehat{\Sigma} \setminus \{x\})^*$. Since $|\Pi \setminus \{x\}| = |\Pi| - 1 = k$, then by the inductive hypothesis, there exists some $w'' \in \Sigma^*$ such that $w' \sim_{D'} w''$. Since $D' \subseteq D$, then $w' \sim_D w''$. Since $\sim_D$ is transitive, then $w \sim_D w''$. Then there exists a $w'' \in \Sigma^*$ such that $w \sim_D w''$.

Since $\Gamma_D(D)$ is acyclic, then by the principle of induction, there exists a $w' \in \Sigma^*$ such that $w \sim_D w'$. □

Lemma B.9. *Let $\Sigma \subseteq \widehat{\Sigma}$ be an alphabet with $\Pi = \widehat{\Sigma} \setminus \Sigma$ and $R \subseteq \widehat{\Sigma}^* \times \widehat{\Sigma}^*$ finite. If $D \subseteq R$ is a family of $\widehat{\Sigma}$-defining relations for Π with $\Gamma_D(D)$ acyclic, then there exists a $Q \subseteq \Sigma^* \times \Sigma^*$ with $|Q| \leq |R \setminus D|$ and a finite sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid Q \cup D \rangle$ of length between k and $2k$ where $k = |R| - |R \cap (\Sigma^* \times \Sigma^*)| - |D|$.*

Proof. Let $S = R \cap (\Sigma^* \times \Sigma^*)$ and $\overline{S} = R \setminus (S \sqcup D)$. Then $R = S \sqcup D \sqcup \overline{S}$. The proof follows by induction on $|\overline{S}|$ in this decomposition.

- **Base Case**. Assume that R decomposes into $S \sqcup D \sqcup \overline{S}$ with $|\overline{S}| = 0$. Then $\langle \widehat{\Sigma} \mid R \rangle = \langle \widehat{\Sigma} \mid S \cup D \rangle$. Then there exists a length 0 sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid S \cup D \rangle$. Clearly $|S| = |R \setminus D|$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if R decomposes as $S \sqcup D \sqcup \overline{S}$ with $|\overline{S}| = k$, then there exists some $Q \subseteq \Sigma^* \times \Sigma^*$ with $|Q| \leq |R \setminus D|$ and a sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid Q \cup D \rangle$ of length at most $2k$.
- **Inductive Step**. Assume that R decomposes into $S \sqcup D \sqcup \overline{S}$ with $|\overline{S}| = k+1$. Let $q \approx w$ in $\overline{S}$. Then by Lemma B.8, there exists $q' \in \Sigma^*$ and $r' \in \Sigma^*$ such that $q \sim_D q'$ and $r \sim_D r'$. Then by Lemma B.7, there exists a finite sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid R' \rangle$ of length between 1 and 2, where $R' = R \cup \{q' \approx w'\} \setminus \{q \approx w\}$. Then R' decomposes as $S' \sqcup D \sqcup \overline{S}'$ where $S' = S \cup \{q' \approx r'\}$ and $\overline{S}' = \overline{S} \setminus \{q \approx r\}$. Since $q \approx_S r$, then $|\overline{S}'| = |\overline{S}| - 1 = k$. Then the inductive hypothesis holds, and there exists some $Q \subseteq \Sigma^* \times \Sigma^*$ with $|Q| \leq |R' \setminus D|$ and a sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R' \rangle$ and $\langle \widehat{\Sigma} \mid Q \cup D \rangle$ of length between k and $2k$. Then there exists a sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid Q \cup D \rangle$ of length between $k+1$ and $2(k+1)$. Since $|S'| \leq |S|$ and $|\overline{S}'| = |\overline{S}|$, then $|Q| \leq |S'| + |\overline{S}'| \leq |S| + |\overline{S}| = |R \setminus D|$.

By the principle of induction, there exists a $Q \subseteq \Sigma^* \times \Sigma^*$ with $|Q| \leq |R \setminus D|$ and a sequence of Tietze transformations between $\langle \widehat{\Sigma} \mid R \rangle$ and $\langle \widehat{\Sigma} \mid Q \cup D \rangle$ of length between $|\overline{S}|$ and $2|\overline{S}|$. □

Lemma B.10. *If D is a family of Σ-defining relations for Π with Π finite and $\Gamma_D(D)$ acyclic, then there exists a relation $x \approx w$ in D such that $D \setminus \{x \approx w\}$ is a family of $(\Sigma \setminus \{x\})$-defining relations for $\Pi \setminus \{x\}$.*

Proof. Since Π is finite and $\Gamma_D(D)$ is acyclic, then by Lemma B.1, there exists some vertex $x \in \Pi$ such that x has no parents in $\Gamma_D(D)$. Since x is a vertex in $\Gamma_D(D)$, then there exists a Σ-defining relation $x \approx w$ in D. Let $D' = D \setminus \{x \approx w\}$. Let $y \approx q$ be a Σ-defining relation in D' with $n = |q|$. Let $k \in [n]$. Assume for

the intent of contradiction that $q_k = x$. Then there exists some $q', q'' \in \Sigma^*$ such that $q = q' \cdot x \cdot q''$. Then x is a child of y by definition. However, x has no parents by assumption. Then $q_k \neq x$ by contradiction. Since k was arbitrary, then $q \in (\Sigma \setminus \{x\})^*$. Since $y \approx q$ was arbitrary, then D' is a family of $(\Sigma \setminus \{x\})$-defining relations for $\Pi \setminus \{x\}$. □

Theorem B.11. *Let $\Sigma' \subseteq \Sigma$ be an alphabet with $\Pi = \Sigma \setminus \Sigma'$ finite and $R \subseteq \Sigma^* \times \Sigma^*$ finite. If $D \subseteq R$ is a family of defining relations for Π with $\Gamma_D(D)$ acyclic, then there exists a $Q \subseteq (\Sigma')^* \times (\Sigma')^*$ with $|Q| \leq |R \setminus D|$ and a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma' \mid Q \rangle$ of length between $n+k$ and $2n+k$ where $n = |R| - |R \cap ((\Sigma')^* \times (\Sigma')^*)| - |D|$ and $k = |\Pi|$.*

Proof. By Lemma B.9, there exists a $Q \subseteq (\Sigma')^* \times (\Sigma')^*$ with $|Q| \leq |R \setminus D|$ and a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid Q \cup D \rangle$ of length between n and $2n$. The proof follows by induction on $|\Pi|$.

- **Base Case**. If $|\Pi| = 0$, then $|D| = 0$. Then $\langle \Sigma \mid Q \cup D \rangle = \langle \Sigma' \mid Q \rangle$. Then there exists a length 0 sequence of Tietze transformations between $\langle \Sigma \mid Q \cup D \rangle$ and $\langle \Sigma' \mid Q \rangle$.
- **Inductive Case**. Assume that for some $k \in \mathbb{N}$, if $|\Pi| = k$ and $\Pi_D(D)$ is acyclic, then there exists a length k sequence of Tietze transformations between $\langle \Sigma \mid Q \cup D \rangle$ and $\langle \Sigma' \mid Q \rangle$.
- **Inductive Step**. Assume that $|\Pi| = k+1$ and $\Gamma_D(D)$ is acyclic. Since Π is finite and $\Gamma_D(D)$ is acyclic, then by Lemma B.1 there exists some vertex $x \in \Pi$ such that x has no parents in $\Gamma_D(D)$. Define $\Lambda = \Sigma \setminus \{x\}$ and $D' = D \setminus \{r_x\}$. Let $y \approx_{D'} w$ and assume for the intent of contradiction that $w \notin \Lambda^*$. Then there exists $w', w'' \in \Sigma^*$ such that $w = w' \cdot x \cdot w''$. Then x is a child of y in $\Gamma_D(D)$. However, x has no parents by assumption. Then $w \in \Lambda^*$ by contradiction. Since $y \approx_{D'} w$ was arbitrary, then $D' \subseteq \Lambda^* \times \Lambda^*$. Then $Q \cup D' \subseteq \Lambda^* \times \Lambda^*$. It follows that $\langle \Sigma \mid Q \cup D \rangle \cong \langle \Lambda \mid Q \cup D' \rangle$ by **Gen**$(-)$. Furthermore, D' is a family of defining relations for $\Pi \setminus \{x\}$. Since $|\Pi \setminus \{x\}| = k$ and $\Gamma_D(D')$ is acyclic by Lemma B.4, then there exists a length k sequence of Tietze transformations between $\langle \Lambda \mid Q \cup D' \rangle$ and $\langle \Sigma' \mid Q \rangle$ by the inductive hypothesis. Then there exists a length $k+1$ sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma' \mid Q \rangle$.

Then by the principle of induction, there exists a length k sequence of Tietze transformations between $\langle \Sigma \mid Q \cup D \rangle$ and $\langle \Sigma' \mid Q \rangle$. Then there exists a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma' \mid Q \rangle$ of length between $n+k$ and $2n+k$. □

B.2.4 The Derived Generator Graph for $W(E_8)$

The derived generators for $W(E_8)$ are $\Pi := \Sigma_D \setminus \{X_0, CX_{0,1}, CCX_{1,2}, K_{1,2}\}$. The defining relations $D \subseteq R_0$ for $\Pi \subseteq \Sigma_D$ are given by Relation (1) through to Relation (19) in Figure 4. An illustration of the derived generator graph $\Gamma_D(D)$ can be found in Figure 9. Since this graph is acyclic, then Theorem B.6 and Theorem B.11 apply. The derived generator graphs for $O(n, \mathbb{D})$ have paths of length at most one, and are therefore trivially acyclic.

B.3 Derivational Proofs and Tietze Transformation

Assume that $G \cong \langle \Sigma \mid R \rangle$ with semantic interpretation $[\![\cdot]\!]_G$. During proofs based on Tietze transformations, it is often necessary to find a sequence of **Rel**$(-)$ and **Rel**$(+)$ transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid Q \rangle$. For example, this case arises in Section 4, where $R = R_{E8} \cup R_{D(E8)}$ and $Q = R_{E8(D)} \cup R_D$. If $[\![\cdot]\!]_G$ induces an isomorphism and every relation $w \approx_Q w'$ satisfies $[\![w]\!]^* = [\![w']\!]$, then using Theorem A.6 there exists a sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid R \rangle$. Eliminating the relations in R

requires more care. For example, if $r \in R$ is not derivable from $(R \cup Q) \setminus \{r\}$, then $\langle \Sigma \mid R \cup Q \rangle$ is a proper quotient of $\langle \Sigma \mid Q \rangle$. Instead, it must be shown that for each $w \approx_R w'$, it follows that $w \sim_Q w'$. One way to approach this problem is to first derive some auxiliary relations A from Q, and then use $Q \cup A$ to derive R. However, transforming these derivations into a sequence of Tietze transformations is often tedious, and not well-aligned with the process of proof discovery. On the other hand, if the derivations are not transformed into a valid sequence of Tietze transformations, then it is possible to obtain invalid proofs, such as those with cyclic derivations (see Example B.16).

This section formalizes the ad-hoc proof technique described above,and identifies sufficient conditions for when such a family of derivations induces a valid sequence of Tietze transformations between $\langle \Sigma \mid R \rangle$ and $\langle \Sigma \mid Q \rangle$. In the following definitions, $L(\Sigma) = \mathbb{N} \times (\Sigma^* \times \Sigma^*)$ will represent a set of indexed relations over Σ. For example, let $(n, (w, w')) \in L(\Sigma)$. The index n in $(n, (w, w'))$ indicates that $(n, (w, w'))$ is a derived relation, and allows for multiple derivations of the same relation. More concretely, if $n = 3$, $w = a \cdot b$, and $w' = x \cdot y \cdot z$, then $(n, (w, w'))$ is the third derivation that yields $a \cdot b \approx x \cdot y \cdot z$.

Definition B.12 (Derivational Proof)**.** A *derivational proof in* $\langle \Sigma \mid R \rangle$ is a subset $P \subseteq L(\Sigma) \times (L(\Sigma) \cup R)^*$ which satisfies the following conditions.

- **Indexed**. For all $(\ell, d) \in P$ and $(\ell', d') \in P$ distinct, $\ell \neq \ell'$.
- **Well-founded**. For all $(\ell, d) \in P$ and $k \in \{1, 2, \ldots, |d|\}$, either $d_k \in R$ or $d_k \in L(\Sigma)$ and there exists some $d' \in (L(\Sigma) \cup R)^*$ such that $(d_k, d') \in P$.
- **Valid**. For all $(\ell, d) \in P$ with $(n, (w, w')) = \ell$ and $m = |d|$, there exists some $v \in (\Sigma^*)^{m+1}$ such that $v_1 = w$, $v_{m+1} = w'$, and for all $k \in [m]$ either $d_k \in R$ and $v_k \xrightarrow{d_k} v_{k+1}$ or $(n', r) = d_k$ and $v_k \xrightarrow{r} v_{k+1}$.

A set $Q \subseteq \Sigma^* \times \Sigma^*$ is *entailed by* P, written $P \models Q$, if $Q \subseteq R \cup \{r \mid ((n, r), d) \in P\}$.

Definition B.13 (Proof Substitution)**.** Let P be a proof in $\langle \Sigma \mid R \rangle$. If $(\ell, d) \in P$, $d \in (L(\Sigma) \cup R)^*$, and $P' = (P \setminus \{(\ell, d)\}) \cup \{(\ell, d')\}$ is a proof for $\langle \Sigma \mid R \rangle$, then we say that P' is a *substitution of* P *by* d' *at* ℓ, written $P[\ell \mapsto d']$.

Definition B.14 (Derivation Graph)**.** Let P be a proof in $\langle \Sigma \mid R \rangle$. The *derivation graph for* P is the digraph $\Gamma_D(P) = (V, E)$ such that $V = \{\ell \mid (\ell, d) \in P\}$ and $(\ell, \ell') \in E$ if and only if there exists $(\ell, d) \in P$ and $k \in \{1, 2, \ldots, |d|\}$ such that $d_k = \ell'$.

Example B.15 (Derivations and Substitutions)**.** Consider $G = \langle x, y \mid x^2 \approx \varepsilon, y^2 \approx \varepsilon \rangle$. It is not hard to show $x \cdot y \cdot x \cdot y^2 \cdot x \cdot y \cdot x \sim \varepsilon$. However, it helps to first prove that $x \cdot y^2 \cdot x \sim \varepsilon$. This can be written as a derivation proof. There will be two derivations, with labels $\ell = (0, (x \cdot y^2 \cdot x, \varepsilon))$ and $\ell' = (0, (x \cdot y \cdot x \cdot y^2 \cdot x \cdot y \cdot x, \varepsilon))$ respectively. Associated with ℓ and ℓ' are two derivations d and d', defined as follows. Let $r = (x^2, \varepsilon)$ and $r' = (y^2, \varepsilon)$.

$$(d) : \; x \cdot y \cdot y \cdot x \xrightarrow{r'} x \cdot x \xrightarrow{r} \varepsilon$$
$$(d') : \; x \cdot y \cdot x \cdot y \cdot y \cdot x \cdot y \cdot x \xrightarrow{\ell} x \cdot y \cdot y \cdot x \xrightarrow{\ell} \varepsilon$$

These pieces can be assembled into a derivational proof $P = \{(\ell, d), (\ell', d')\}$. This proof is indexed, since ℓ and ℓ' each appear exactly once as labels. This proof is well-formed, since r, r', d, and d' are all in-scope. The proof is valid, since each step of each derivation follows. Of course, it is possible to expand out d' using the steps of d. This yields a new derivation d'' defined as follows.

$$(d'') : \; x \cdot y \cdot x \cdot y \cdot y \cdot x \cdot y \cdot x \xrightarrow{r'} x \cdot y \cdot x \cdot x \cdot y \cdot x \xrightarrow{r} x \cdot y \cdot y \cdot x \xrightarrow{r'} x \cdot x \xrightarrow{r} \varepsilon$$

This new derivational proof corresponds to the substitution $P[\ell' \mapsto d'']$. In this new proof, the derivation of ℓ' no longer relies on the lemma ℓ. Later in this section, lemma eliminating substitutions will be used to extract Tietze transformations from derivational proofs with acyclic derivation graphs. □

Example B.16 (Cyclic Derivations)**.** Consider $G = \langle x, y \mid x^2 \approx \varepsilon, y^2 \approx \varepsilon \rangle$. It is not hard to show that $G \cong \mathbb{Z}_2 \star \mathbb{Z}_2$, where $(\star)$ denotes the free product of groups. It follows that G is non-abelian. In particular, $\pi_G(x \cdot y) \neq \pi_G(y \cdot x)$. Now, consider the proof $P = \{((0,(x \cdot y, y \cdot x)), d), ((0,(x \cdot y^3, y \cdot x \cdot y)), d')\}$, where d and d' are defined as follows. Let $r = (y^2, \varepsilon)$, $\ell = (0,(x \cdot y, y \cdot x))$ and $\ell' = (0,(x \cdot y^3, y \cdot x \cdot y^2))$.

$$(d) : x \cdot y \xleftarrow{r} x \cdot y^3 \xrightarrow{\ell'} y \cdot x \cdot y^2 \xrightarrow{r} y \cdot x$$
$$(d') : x \cdot y^3 \xrightarrow{\ell} y \cdot x \cdot y^2$$

This proof is indexed, well-formed, and valid. However, these derivations suggest that $x \cdot y \sim y \cdot x$. The problem in this proof is that d depends on d' and d' depends on d. In other words, the proof is self-referential. It will be shown later in this section that if $\Gamma_D(P)$ is acyclic, then P is not self-referential. This motivates the requirement that $\Gamma_D(P)$ is acyclic throughout the rest of this section. □

B.3.1 Substitutions and Derivation Graphs

This section relates derivational proofs and substitutions to the derivation graphs they induce. When a derivation (ℓ, d) depends on a derivation (ℓ', d'), we say that ℓ' is a lemma for ℓ. In Lemma B.17, lemma-free derivations are characterized by their vertices in a derivation graph. Similarly, Lemma B.18 characterizes lemma-free proofs in terms of their derivation graphs. As expected, Lemma B.19 shows that if a substitution only eliminates the use of lemmas, such as in Example B.15, then the resulting derivation graph is a subgraph of the original graph. To this end, Lemma B.20 provides sufficient conditions for a valid substitution. Together, these four lemmas give a graph-theoretic characterization of the lemma substitution in Example B.15.

Lemma B.17. *Let P be a proof for $\langle \Sigma \mid R \rangle$. If $(\ell, d) \in P$ and ℓ has no children in $\Gamma_D(P)$, then $d \in R^*$.*

Proof. Let $k \in \{1, 2, \ldots, |d|\}$. Assume for the intent of contradiction that $d_k \notin R^*$. Then (ℓ, d_k) is an edge in $\Gamma_D(P)$. However, ℓ has no children in $\Gamma_D(P)$. Then $d_k \in R$. Since k was arbitrary, then $d \in R^*$. □

Lemma B.18. *If P is a proof for $\langle \Sigma \mid R \rangle$, then $\Gamma_D(P)$ is edgeless if and only if $d \in R^*$ for all $(\ell, d) \in P$.*

Proof. Let $\Gamma_D(P) = (V, E)$. Consider the contrapositive statement. Then there exists an $(\ell, \ell') \in E$. This is true if and only if there exists a $(\ell, d) \in P$ and $k \in \{1, 2, \ldots, |d|\}$ such that $d_k = \ell'$. This is true if and only if there exists an $(\ell, d) \in P$ such that $d \notin R^*$. □

Lemma B.19. *Let P be a proof for $\langle \Sigma \mid R \rangle$ and $(V, E) = \Gamma_D(P)$. If $(\ell, d) \in P$ and there exists some $d' \in R^*$ such that $P' = P[l \mapsto d']$ is also a proof for $\langle \Sigma \mid R \rangle$, then $\Gamma_D(P') = (V, E_\ell)$ where $E_\ell = \{(\ell', \ell'') \in E \mid \ell \neq \ell'\}$.*

Proof. Let $\Gamma_D(P') = (V', E')$.

- **Vertices ($\subseteq$).** Let $\ell' \in V'$. Then there exists some $\delta \in (L(\Sigma) \cup R)^*$ such that $(\ell', \delta) \in P'$. Then either $(\ell', \delta) \in P$ or $(\ell', \delta) = (\ell, d')$. If $(\ell', \delta) \in P$, then $\ell \in V$. If $(\ell', \delta) = (\ell, d')$, then $\ell' \in V$ since $(\ell, d) \in P$. In either case $\ell' \in V$. Since ℓ' was arbitrary, then $V' \subseteq V$.
- **Vertices ($\supseteq$).** Let $\ell' \in V$. Then there exists some $\delta \in (L(\Sigma) \cup R)^*$ such that $(\ell', \delta) \in P$. Then either $\ell' = \ell$ or $\ell' \neq \ell$. If $\ell' = \ell$, then $\ell' \in V'$ since $(\ell', d') \in P'$. If $\ell' \neq \ell$, then $\ell' \in V'$ since $(\ell', \delta) \in P \setminus \{(\ell, d)\}$. In either case, $\ell' \in V$, Since ℓ' was arbitrary, then $V' \subseteq V$.

– **Edges ($\subseteq$).** Let $(\ell', \ell'') \in E'$. Then there exists some $\delta \in (L(\Sigma) \cup R)^*$ and $k \in \{1, 2, \ldots, |\delta|\}$ such that $(\ell', \delta) \in P'$ and $\delta_k = \ell''$. Then $\delta \notin R^*$. Consequently, $\delta \neq d'$. Then $\ell \neq \ell'$, since P' is indexed. Then $(\ell', \delta) \in P$. Consequently, $(\ell', \ell'') \in E$. Since $\ell' \neq \ell$, then $(\ell', \ell'') \in E_\ell$. Since (ℓ', ℓ'') was arbitrary, then $E' \subseteq E_\ell$.

– **Edges ($\supseteq$).** Assume that $(\ell', \ell'') \in E_\ell$. Then $\ell' \neq \ell$ and $(\ell', \ell'') \in E$. Then there exists some $\delta \in (L(\Sigma) \cup R)^*$ and $k \in \{1, 2, \ldots, |\delta|\}$ such that $(\ell', \delta) \in P$ and $\delta_k = v$. Since $\ell' \neq \ell$, then $(\ell', \delta) \in P'$. Then $(\ell', \ell'') \in E'$. Since (ℓ', ℓ'') was arbitrary, then $E_\ell \subseteq E'$.

Then $\Gamma_D(P') = (V, E_\ell)$. □

Lemma B.20. *Let P be a proof for $\langle \Sigma \mid R \rangle$ and $V = \{\ell \mid (\ell, d) \in P\}$. If $(\ell, d) \in P$, $d' \in (V \cup R)^*$, and (ℓ, d') is valid, then $P[\ell \mapsto d']$ is a proof for $\langle \Sigma \mid R \rangle$.*

Proof. Let $P' = [\ell \mapsto d']$. It must be shown that P' is indexed, well-formed, and valid.

– **Indexed**. Let $(\ell', \delta) \in P'$ and $(\ell'', \delta') \in P'$ distinct. Without loss of generality, assume $\ell' \neq \ell$. Then $(\ell', \delta) \in P$. Now, these are two cases to consider, depending on whether $\ell'' = \ell$. If $\ell'' = \ell$, then $(\ell'', d) \in P$ and $\ell' \neq \ell''$ since P is indexed. If $\ell'' \neq \ell$, then $(\ell'', \delta') \in P$ and $\ell' \neq \ell''$ since P is indexed. In either case $\ell' \neq \ell'$. Since (ℓ', δ) and (ℓ'', δ') were arbitrary, then P' is indexed.

– **Well-formed**. Let $(\ell', \delta) \in P'$. There are two cases to consider, depending on whether $\ell' = \ell$.
  – If $\ell' = \ell$, then $\delta = d'$, since P' is indexed. Then $\delta \in (V \cup R)^*$. Then for all $k \in \{1, 2, \ldots, |\delta|\}$, either $\delta_k \in R$ or $\delta_k \in V$. If $\delta_k \in V$ and $\delta_k = \ell$, then $(\delta_k, d') \in P'$. If $\delta_k \in V$ and $\delta_k \neq \ell$, then there exists some $\delta' \in (L(\Sigma) \cup R)^*$ such that $(d_k, \delta') \in P$. Since $\delta_k \neq \ell$, then $(\delta_k, \delta') \in P'$. In either case, if $\delta_k \in V$, then there exists some $\delta' \in (L(\Sigma) \cup R)^*$ such that $(d_k, d'') \in P'$ Since k was arbitrary, then (ℓ', δ) is well-formed.
  – If $\ell' \neq \ell$, then $(\ell', \delta) \in P$. Let $k \in \{1, 2, \ldots, |\delta|\}$. Since P is well-formed, then either $\delta_k \in R$ or there exists some $\delta' \in (L(\Sigma) \cup R)^*$ such that $(\delta_k, \delta') \in P$. There are two cases to consider, depending on whether $\delta_k = \ell$. If $\delta_k = \ell$, then $(\ell, d') \in P'$. If $\delta_k \neq \ell$, then $(\delta_k, \delta') \in P'$. In either case, there exists some $\delta' \in (L(\Sigma) \cup R)^*$ such that $(\delta_k, \delta') \in P'$. Since k was arbitrary, then (ℓ', δ) is well-formed.

  In either case, (ℓ', δ) is well-formed. Since (ℓ', δ) was arbitrary, then P' is well-formed.

– **Valid**. Let $(\ell', \delta) \in P'$. Now these are two cases to consider, depending on whether $\ell' = \ell$. If $\ell' = \ell$, then $\delta = d'$ since P' is indexed, and consequently δ is valid by assumption. If $\ell' \neq \ell$, then $(\ell', \delta) \in P$ and δ is valid by the validity of P. In either case, (ℓ', δ) is valid. Since (ℓ', δ) was arbitrary, then P' is valid.

Then P' is a proof for $\langle \Sigma \mid R \rangle$. □

B.3.2 From Derivational Proofs to Tietze Transformations

Let $\langle \Sigma \mid R \rangle$ be a monoid presentation. The goal of this section is to prove Theorem B.25, which states that acyclic derivational proofs are sound for the isomorphism of finite monoid presentations. The completeness of acyclic derivational proofs follows immediately from the fact that every statement of the form $w \sim_R w'$ corresponds to at least one finite derivation. The proof proceeds as follows. First, Lemma B.21 shows that the lemma substitutions outlined in Example B.15 preserve the structure of derivational proofs. This is used in Lemma B.22 to show that if a derivation (ℓ, d) in a proof P depends on a lemma (ℓ', d') which follows directly from R, then d can be rewritten so that all dependencies on ℓ'

are removed without introducing any new dependencies. This is used repeated in Lemma B.23 to show that any derivation (ℓ, d) which depends only on lemmas which follow directly from R, can be rewritten to also follow directly from R. This is extended in Lemma B.24 to show that any finite and acyclic derivational proof P can be written into a proof P' such that every derivation depends only on R. Then each derivation in P' follows from $(\sim_R)$, and Theorem B.25 follows immediately by induction.

Lemma B.21. *Let P be a proof for $\langle \Sigma \mid R \rangle$. If (ℓ, d) and (ℓ', d') are derivations in P and there exists $\delta, \delta' \in (L(\Sigma) \cup R)^*$ such that $d = \delta \cdot \ell' \cdot \delta'$, then $P[\ell \mapsto \delta \cdot d' \cdot \delta']$ is a proof for $\langle \Sigma \mid R \rangle$.*

Proof. Let $L_P = \{\ell \mid (\ell, d) \in P\}$, $d'' = \delta \cdot d' \cdot \delta'$, and $(x, (q, r)) = \ell$. Since $\delta, d', \delta' \in (L_p \cup R)^*$, then $d'' \in (L_p \cup R)^*$. Since P is valid, then there exists $v \in (\Sigma^*)^{n+1}$ such that $v_1 = q$, $v_{n+1} = r$, and $v_k \xrightarrow{d_k} v_{k+1}$ for all $k \in [n]$, where $n = |d|$. Let $m = |\delta|$. Since P is valid and $v_m \xrightarrow{\ell'} v_{m+1}$, then there exists some $u \in (\Sigma^*)^{s+1}$ such that $u_1 = v_m$, $u_{s+1} = v_{m+1}$, and $u_k \xrightarrow{d'_k} u_{k+1}$ for all $k \in [s]$, where $s = |d'|$. Define $v' = (v_1, \ldots, v_m, u_1, \ldots, u_s, v_{m+1}, \ldots, v_n)$. Let $k \in \{1, 2, \ldots, |d''|\}$. There are five cases to consider.

- If $k < m$, then $v'_k = v_k$, $v'_{k+1} = v_{k+1}$, and $d''_k = \delta_k = d_k$. Then $v'_k \xrightarrow{d''_k} v'_{k+1}$.
- If $k = m$, then $v'_k = v_m$, $v'_{k+1} = u_1 = v_{m+1}$, and $d''_k = \delta_k$. Then $v'_k \xrightarrow{d''_k} v'_{k+1}$.
- If $m < k < m + s$, then $v'_k = u_{k-m}$, $v'_{k+1} = u_{k-m+1}$, and $d''_k = d'_{k-m}$. Then $v'_k \xrightarrow{d''_k} v'_{k+1}$.
- If $k = m + s$, then $v'_k = u_s$, $v'_{k+1} = v_{m+1} = u_{s+1}$, and $d''_k = d'_s$. Then $v'_k \xrightarrow{d''_k} v'_{k+1}$.
- If $k > m + s$, then $v'_k = u_{k-s}$, $v'_{k+1} = u_{k+1-s}$, and $d''_k = \delta'_{k-m-s}$. Then $v'_k \xrightarrow{d''_k} v'_{k+1}$.

In each case, $v'_k \xrightarrow{d''_k} v'_{k+1}$. Since k was arbitrary, then (ℓ, d'') is valid. Then $P[\ell \mapsto d'']$ is a proof by Lemma B.20. □

Lemma B.22. *Let P be a proof for $\langle \Sigma \mid R \rangle$. If (ℓ, ℓ') is a maximal path rooted at ℓ in $\Gamma_D(P)$, then there exists a $\hat{d} \in (L(\Sigma) \cup R)^*$ such that $P' = P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$ and ℓ has one less child in $\Gamma_D(P')$.*

Proof. Let $f : (L(\Sigma) \cup R)^* \to \mathbb{N}$ count the number of occurrences of ℓ' in a derivation. Given a proof Q for $\langle \Sigma \mid R \rangle$, let $C_Q : L(\Sigma) \to \mathscr{P}(L)$ map each $\ell \in L(\Sigma)$ to its children in $\Gamma_D(Q)$. Since (ℓ, ℓ') is a maximal path, then ℓ' has no children in $\Gamma_D(P)$. The proof follows by induction on the number of occurrences of ℓ' in the derivation.

- **Base Case**. Let $\hat{d} \in (L(\Sigma) \cup R)^*$ with $P' = P[\ell \mapsto \hat{d}]$ a proof for $\langle \Sigma \mid R \rangle$ and $C_P(\ell) = C_{P'}(\ell) \cup \{\ell'\}$. Assume that $f(\hat{d}) = 0$. Then $\hat{d}_j \neq \ell'$ for all $k \in [n]$, where $n = |\hat{d}|$. Then $(\ell, \ell') \notin \Gamma_D(P')$. Then $\ell' \notin C_{P'}(\ell)$. Then $C_{P'}(\ell) = C_P(\ell) \setminus \{\ell'\}$. Since $\ell' \in C_P(\ell)$, then $|C_{P'}(\ell)| = |C_P(\ell)| - 1$.
- **Inductive Hypothesis**. Let $d' \in (L(\Sigma) \cup R)^*$ such that $P' = P[\ell \mapsto d']$ a proof for $\langle \Sigma \mid R \rangle$ and $C_P(\ell) = C_{P'}(\ell) \cup \{\ell'\}$. Assume that for some $k \in \mathbb{N}$, if $f(d') = k$, then exists a $\hat{d} \in (L(\Sigma) \cup R)^*$ such that $P'' = P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$ and $|C_{P''}(\ell)| = |C_P(\ell)| - 1$.
- **Inductive Step**. Under the conditions of the inductive hypothesis, assume that $f(d') = k + 1$. Then there exists some $\delta, \delta' \in (L(\Sigma) \cup R)^*$ such that $d' = \delta \cdot \ell' \cdot \delta'$. Since ℓ' is a vertex in $\Gamma_D(P)$, then there exists some $d'' \in (L(\Sigma) \cup R)^*$ such that $(\ell, d'') \in P$. Define $\hat{d} = \delta \cdot d'' \cdot \delta'$. By Lemma B.21, $P'' = P[l \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$. Let $\ell'' \in C_P(\ell) \setminus \{\ell'\}$. Then there exists some $k \in \{1, 2, \ldots, |\delta|\}$ such that $d_k = \ell''$. Since $\ell'' \neq \ell$, then without loss of generality $k \leq |\delta|$ and $\hat{d}_k = \delta_k = \ell''$. Then $\ell'' \in C_{P''}(\ell)$. Since ℓ'' was arbitrary, then $C_P(\ell) \setminus \{\ell'\} \subseteq C_{P''}(\ell)$. Next, let $\ell'' \in C_{P''}(\ell)$. Then there

exists some $k \in \{1,2,\ldots,|\hat{d}|\}$ such that $\hat{d}_k = \ell''$. Since $d \in (R^*)$, then without loss of generality $k \leq |\delta|$ and $\delta_k = \hat{h}_k = \ell''$. Then $\ell'' \in C_P(\ell)$. Since ℓ'' was arbitrary, then $C_{P''}(\ell) \subseteq C_P(\ell)$. Then $C_P(\ell) = C_{P''}(\ell) \cup \{\ell'\}$. Since $k+1 = f(d') = f(\delta) + 1 + f(\delta')$, then $f(\hat{d}) = f(\delta) + f(\delta') = k$. Then by the inductive hypothesis, there exists some $\hat{d'} \in (L(\Sigma) \cup R)^*$ such that $Q = P'[\ell \mapsto \hat{d'}]$ is a proof for $\langle \Sigma \mid R \rangle$ and $|C_Q(\ell)| = |C_{P'}(\ell)| - 1 = |C_P(\ell)| - 1$. Since $Q = P[\ell \mapsto \hat{d'}]$ by definition, then the inductive step holds.

It follows by definition that $d \in (L(\Sigma) \cup R)^*$ and $P = P[\ell \mapsto d]$. Then by the principle of induction, there exists a $\hat{d} \in (L(\Sigma) \cup R)^*$ such that $P' = P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$ and $|C_{P''}(\ell)| = |C_P(\ell)| - 1$. □

Lemma B.23. *Let P be a proof for $\langle \Sigma \mid R \rangle$. If $(\ell, d) \in P$ and all paths rooted at ℓ in $\Gamma_D(P)$ have length at most one, then there exists a $\hat{d} \in R^*$ such that $P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$.*

Proof. Let $\Gamma_D(P) = (V, E)$. Since $|V| = |P|$ and all paths rooted at ℓ have length at most one, then the number of paths rooted at ℓ in $\Gamma_D(P)$ is finite. The proof follows by induction on the number of paths rooted at ℓ in $\Gamma_D(P)$.

- **Base Case**. Assume that $\Gamma_D(P)$ has zero paths rooted at ℓ. Then $d \in R^*$ by Lemma B.17. Then $P = P[l \mapsto d]$ with $d \in R^*$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $\Gamma_D(P)$ has k paths rooted at ℓ, then there exists a $\hat{d} \in R^*$ such that $P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$.
- **Inductive Step**. Assume that $\Gamma_D(P)$ has $k+1$ paths rooted at ℓ. Then there exists at least one path rooted at ℓ in $\Gamma_D(P)$. Since all paths rooted at ℓ in $\Gamma_D(P)$ have length one, then there exists some path (ℓ, ℓ') in $\Gamma_D(P)$ such that ℓ' has no children in $\Gamma_D(P)$. Then by Lemma B.22, there exists some $d' \in (L(\Sigma) \cup R)^*$ such that $P' = P[\ell \mapsto d']$ is a proof for $\langle \Sigma \mid R \rangle$ and ℓ has k children in $\Gamma_D(P')$. Since all paths rooted at ℓ have length one, then $\Gamma_D(P')$ has k paths rooted at ℓ. Then by the inductive hypothesis, then there exists a $\hat{d} \in R^*$ such that $P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$. Then the inductive step holds.

Then by the principle of induction, there exists a $\hat{d} \in R^*$ such that $P[\ell \mapsto \hat{d}]$ is a proof for $\langle \Sigma \mid R \rangle$. □

Lemma B.24. *Let P be a finite proof for $\langle \Sigma \mid R' \rangle$ with $R' \subseteq R \subseteq \Sigma^* \times \Sigma^*$. If $P \models R$ and $\Gamma_D(P)$ is acyclic, then there exists a proof P' for $\langle \Sigma \mid R' \rangle$ such that $|P'| = |P|$, $P' \models R$, and $d \in (R')^*$ for all $(\ell, d) \in P'$.*

Proof. Let $f : L(\Sigma) \times (L(\Sigma) \cup R')^* \to \mathbb{N}$ count the number of vertices with children in the derivation graph of a proof. Since $\Gamma_D(P)$ has $|P|$ vertices, then $f(P) \leq |P|$. Since P is finite, then $f(P)$ is also finite. The proof follows by induction on $f(P)$.

- **Base Case**. Assume that $f(P) = 0$. Then there are no edges in $\Gamma_D(P)$. Then by Lemma B.18, $d \in (R')^*$ for all $(\ell, d) \in P$.
- **Inductive Hypothesis** Let Q be a proof for $\langle \Sigma \mid R' \rangle$. Assume that for some $k \in \mathbb{N}$, if $f(Q) = k$ with $P \models R$ and $\Gamma_D(Q)$ is acyclic, then there exists a proof P' for $\langle \Sigma \mid R \rangle$ such that $|P'| = |Q|$, $P' \models R$, and $d \in (R')^*$ for all $(\ell, d) \in P'$.
- **Inductive Step**. Let Q be a proof for $\langle \Sigma \mid R' \rangle$. Assume that $f(Q) = k+1$ with $Q \models R$ and $\Gamma_D(Q)$ acyclic. Since $f(Q) > 0$, then there exists at least one edge (ℓ, ℓ') in $\Gamma_D(Q)$. By Lemma B.1, there exists some path $(\ell_0, \ldots, \ell_n)$ in $\Gamma_D(Q)$ such that $\ell_0 = \ell$ every path rooted at ℓ has length at most n. Since (ℓ, ℓ') is a path of length one in $\Gamma_D(Q)$ rooted at ℓ, then $n \geq 1$. Assume for the intent of contradiction that there exists a path of length at least 2 rooted at ℓ_{n-1}. Then there exists a

path (ℓ_{n-1}, x, y) in $\Gamma_D(Q)$. Then $(\ell_0, \ldots, \ell_{n-1}, x, y)$ is a path of length $n+1$ in $\Gamma_D(Q)$ rooted at ℓ. However, all paths rooted at ℓ have length at most n. Therefore, all paths rooted at ℓ_{n-1} have length at most one. Since ℓ_{n-1} is a vertex in $\Gamma_D(Q)$, then there exists some $d \in (L(\Sigma) \cup R')^*$ such that $(\ell_{n-1}, d) \in Q$. Then by Lemma B.23, there exists some $d' \in (R')^*$ such that $P' = Q[\ell_{n-1} \mapsto d']$ is a proof for $\langle \Sigma \mid R' \rangle$. Let $(z_{n-1}, r_{n-1}) = \ell_{n-1}$. Since $Q \models R$ with respect to $\langle \Sigma \mid R' \rangle$, it follows that $R \subseteq R' \cup \{r \mid ((m,r),d) \in P\}$. Then,

$$R \subseteq R' \cup \{r \mid ((z,r),d) \in P \setminus \{(\ell_{n-1}, d)\} \cup \{r_{n-1}\} \subseteq R' \cup \{r \mid ((m,r),d) \in P'\}.$$

Then $P' \models R$ with respect to $\langle \Sigma \mid R' \rangle$. By Lemma B.19, $\Gamma_D(P')$ is also a subgraph of $\Gamma_D(Q)$ with $f(P') = f(Q) - 1 = k$. Since $\Gamma_D(P')$ is a subgraph of $\Gamma_D(Q)$ with $\Gamma_D(Q)$ acyclic, then $\Gamma_D(P')$ is also acyclic. Then by the inductive hypothesis, there exists a proof P'' for $\langle \Sigma \mid R \rangle$ such that $|P''| = |P'|$, $P'' \models R$, and $d \in (R')^*$ for all $(\ell, d) \in P''$. Then $|P''| = (|Q| - 1) + 1 = |Q|$, since Q is indexed. Then the inductive step holds.

Then by the principle of induction, there exists a proof P' for $\langle \Sigma \mid R \rangle$ such that $|P'| = |P|$, $P' \models R$, and $d \in (R')^*$ for all $(\ell, d) \in P$. □

Theorem B.25. *Let P be a finite proof for $\langle \Sigma \mid R' \rangle$ with $R' \subseteq R \subseteq \Sigma^* \times \Sigma^*$. If $P \models R$ and $\Gamma_D(P)$ is acyclic, then there exists a length $|R \setminus R'|$ sequence of **Rel(+)** transformations between $\langle \Sigma \mid R' \rangle$ and $\langle \Sigma \mid R \rangle$.*

Proof. By Lemma B.24, there exists a proof P' for $\langle \Sigma \mid R' \rangle$ with $P' \models R$ and $d \in (R')^*$ for all $(\ell, d) \in P$. The proof follows by induction on $|R \setminus R'|$.

- **Base Case**. If $|R \setminus R'| = 0$, then $R \subseteq R' \subseteq R$. Then $R = R'$ and $\langle \Sigma \mid R \rangle = \langle \Sigma \mid R' \rangle$. Then there exists a length zero sequence of **Rel**(+) transformations between $\langle \Sigma \mid R' \rangle$ and $\langle \Sigma \mid R \rangle$.
- **Inductive Hypothesis**. Let P' be a proof with respect to $\langle \Sigma \mid Q \rangle$ with $Q \subseteq R$ and $P' \models R$. Assume that for some $k \in \mathbb{N}$, if $|R \setminus R'| = k$, then there exists a length k sequence of **Rel**(+) transformations between $\langle \Sigma \mid Q \rangle$ and $\langle \Sigma \mid R \rangle$.
- **Inductive Step**. Under the conditions of the inductive hypothesis, assume that $|R \setminus R'| = k+1$. Then there exists some $r \in R \setminus R'$, say $(w, w') = r$. Since $P' \models R$ and $r \notin R'$, then there exists some $x \in \mathbb{N}$ and $d \in (L(\Sigma) \cup R')^*$ such that $((x,r),d) \in P'$. Since $d \in (R')^*$, then $w \sim_{R'} w'$ by validity of P'. Let $Q = R' \cup \{r\}$. Then $\langle \Sigma \mid R' \rangle \cong \langle \Sigma \mid Q \rangle$ by **Rel**(+). Since $R' \subseteq Q$ and $P' \models R$ with respect to $\langle \Sigma \mid R' \rangle$, then
$$R \subseteq R' \cup \{r \mid ((n,r),d) \in P'\} \subseteq Q \cup \{r \mid ((n,r),d) \in P'\}.$$
Then $P' \models R$ with respect to $\langle \Sigma \mid Q \rangle$. Since $r \in R \setminus R'$, then $|R \setminus Q| = |R \setminus R'| - 1 = k$. Then by the inductive hypothesis, there exists a length k sequence of **Rel**(+) transformations between $\langle \Sigma \mid Q \rangle$ and $\langle \Sigma \mid R \rangle$. Then there exists a length $k+1$ sequence of **Rel**(+) transformations between $\langle \Sigma \mid R' \rangle$ and $\langle \Sigma \mid R \rangle$. Then the inductive step holds.

Then by the principle of induction, there exists a length $|R \setminus R'|$ sequence of **Rel**(+) transformations between $\langle \Sigma \mid R' \rangle$ and $\langle \Sigma \mid R \rangle$. □

C Circuit Decompositions of Coxeter Generators

In Section 4.1, the Coxeter generator r_3 was decomposed into a circuit over Σ_D. In this section, the remaining 7 Coxeter generators are decomposed into circuits over Σ_D. Scalar multiples of the normal vectors are used freely. Recall that $CCX_{0,1}$ is a reflection about the normal vector $|\hat{b}\rangle = |1\rangle \otimes |1\rangle \otimes |-\rangle$. Similarly, $CX_{1,2}$ is a reflection about the normal vector $|\bar{b}\rangle = |-\rangle \otimes |1\rangle \otimes |1\rangle$.

r_1. This generator is defined by the normal vector $|b_1\rangle = |0\rangle \otimes |0\rangle \otimes |-\rangle$. Since $(X_0 \circ X_1)|\hat{b}\rangle = |b_1\rangle$ with $(X_0 \circ X_1)^{-1} = X_1 \circ X_0$, then $r_1 = X_0 \circ X_1 \circ CCX_{0,1} \circ X_1 \circ X_0$.

r_2. This generator is defined by the normal vector $|b_2\rangle = |0\rangle \otimes (|1\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle)/\sqrt{2}$. Since $|b_2\rangle$ and $-|b_2\rangle$ define the same hyperplane, then $-|b_2\rangle$ also defines the same generator. Recall that r_3 is a reflection about the normal vector $|b_3\rangle = |0\rangle \otimes |1\rangle \otimes |-\rangle$. Then $(CX_{2,1})|b_3\rangle = -|b_2\rangle$. Since $CX_{2,1}$ is self-inverse, then $r_2 = CX_{2,1} \circ r_3 \circ CX_{2,1}$. Since $r_3 = X_0 \circ CCX_{0,1} \circ X_0$ with X_0 and $CX_{2,1}$ commuting, then $r_2 = X_0 \circ CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1} \circ X_0$.

r_4. This generator is defined by the normal vector $|b_4\rangle = (|0\rangle \otimes |1\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle \otimes |0\rangle)/\sqrt{2}$. Since $|\bar{b}\rangle = (|0\rangle \otimes |1\rangle \otimes |1\rangle - |1\rangle \otimes |1\rangle \otimes |1\rangle)/\sqrt{2}$, then $(CX_{0,1} \circ CX_{0,2})|\bar{b}\rangle = |b_4\rangle$. Furthermore, since $(CX_{0,1} \circ CX_{0,2})^{-1} = CX_{0,2} \circ CX_{0,1}$, then $r_4 = CX_{0,1} \circ CX_{0,2} \circ CCX_{1,2} \circ CX_{0,2} \circ CX_{0,1}$.

r_5. This generator is defined by the normal vector $|b_5\rangle = |1\rangle \otimes |0\rangle \otimes |-\rangle$. Since $(X_1)|\hat{b}\rangle = |b_5\rangle$ with X_1 self-inverse, then $r_5 = X_1 \circ CCX_{0,1} \circ X_1$.

r_6. This generator is defined by the normal vector $|b_6\rangle = |1\rangle \otimes (|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)/\sqrt{2}$. Since $(CX_{2,1})|\hat{b}\rangle = |b_6\rangle$ with $CX_{2,1}$ self-inverse, then $r_6 = CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1}$.

r_7. This generator is defined by the normal vector $|b_7\rangle = |1\rangle \otimes (|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle)/\sqrt{2}$. Recall that r_6 is a reflection about the normal vector $|b_6\rangle = |1\rangle \otimes (|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)/\sqrt{2}$. It follows that $(CZ_{0,1})|b_6\rangle = |b_7\rangle$. Since $CZ_{0,1}$ is self-inverse, then $r_7 = CZ_{0,1} \circ r_6 \circ CZ_{0,1}$. Furthermore, since $r_6 = CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1}$, then $r_7 = CZ_{0,1} \circ CX_{2,1} \circ CCX_{0,1} \circ CX_{2,1} \circ CZ_{0,1}$.

r_8. The generator is defined by the normal vector $|b_8\rangle = |+\rangle \otimes |+\rangle \otimes |+\rangle$. First, define the operator $M = K_{1,2} \circ X_1 \circ X_2 \circ CZ_{0,2}$. Clearly, $M^{-1} = CZ_{0,2} \circ X_2 \circ X_1 \circ K_{1,2}$. Furthermore,

$$M|\bar{b}\rangle = (K_{1,2} \circ X_1 \circ X_2)(|+\rangle \otimes |1\rangle \otimes |1\rangle) = K_{1,2}(|+\rangle \otimes |0\rangle \otimes |0\rangle) = |b_8\rangle.$$

Therefore, $r_8 = K_{1,2} \circ X_1 \circ X_2 \circ CZ_{0,2} \circ CCX_{1,2} \circ CZ_{0,2} \circ X_2 \circ X_1 \circ K_{1,2}$.

This establishes all decompositions of the Σ_{E8} in terms of Σ_D.

D Constructing the Generators for $W(E_8)$

The section walks through the construction of X_0, $CX_{0,1}$, $CCX_{1,2}$, and $K_{1,2}$ using the Coxeter generators for $W(E_8)$. As suggested in Section 4.1, this construction begins by deriving several diagonal matrices over (± 1).

$$w_1 = r_6 \cdot r_7 \qquad w_2 = r_6 \cdot r_5 \cdot w_1 \cdot r_5 \cdot r_6 \qquad w_3 = r_5 \cdot r_4 \cdot w_2 \cdot r_4 \cdot r_5$$
$$w_4 = r_4 \cdot r_3 \cdot w_3 \cdot r_3 \cdot r_4 \qquad w_5 = r_3 \cdot r_2 \cdot w_4 \cdot r_2 \cdot r_3 \qquad w_6 = r_2 \cdot r_1 \cdot w_5 \cdot r_1 \cdot r_2$$

For example, $[\![w_1]\!]^*_{E8} = CZ_{0,1} \circ CZ_{0,2}$. It is then possible to derive $CCX_{0,1}$ and X_2.

$$w_7 = r_7 \cdot r_8 \cdot r_6 \cdot w_6 \cdot w_4 \cdot w_2 \cdot r_8 \cdot w_6 \cdot w_4 \cdot w_2 \cdot r_6 \cdot r_8 \cdot r_7 \qquad w_8 = r_1 \cdot r_3 \cdot r_5 \cdot w_7$$

Then $[\![w_7]\!]^*_{E8} = CCX_{0,1}$ and $[\![w_8]\!]^*_{E8} = X_2$. Using $CCX_{0,1}$, it is then possible to derive $K_{1,2}$.

$$w_9 = r_6 \cdot w_7 \cdot w_1 \cdot w_7 \cdot r_6 \qquad w_{10} = r_2 \cdot r_6 \cdot w_5 \cdot w_3 \cdot w_2 \cdot r_8 \cdot w_9 \cdot r_8 \cdot w_5 \cdot w_3 \cdot w_2 \cdot w_9$$

Then $[\![w_9]\!]^*_{E8}$ is a diagonal matrix over (± 1) and $[\![w_{10}]\!]^*_{E8} = K_{1,2}$. Next, the permutations are derived.

$$w_{11} = w_{10} \cdot r_4 \cdot w_8 \cdot r_4 \cdot w_{10} \cdot w_8 \qquad w_{12} = r_2 \cdot r_6 \qquad w_{13} = w_{11} \cdot w_{12} \cdot w_{11}$$

It can be validated that $[\![w_{11}]\!]^*_{E8} = \sigma_{1,2}$, $[\![w_{12}]\!]^*_{E8} = \sigma_{0,1}$, and $[\![w_{13}]\!]^*_{E8} = \sigma_{0,2}$. As an immediate consequence, $[\![w_{13} \cdot w_7 \cdot w_{13}]\!]^*_{E8} = CCX_{1,2}$ and $[\![w_{13} \cdot w_8 \cdot w_{13}]\!]^*_{E8} = X_0$. Then by three applications of **Gen**$(+)$, the generators $K_{1,2}$, $CCX_{1,2}$, and X_0 are introduced, alongside the following relations.

$$K_{1,2} \approx w_{10} \qquad CCX_{1,2} \approx w_{13} \cdot w_7 \cdot w_{13} \qquad X_0 \approx w_{13} \cdot w_8 \cdot w_{13}$$

Next, define $w_{14} = w_{12} \cdot X_0 \cdot w_{10} \cdot X_0 \cdot w_{12}$. It can be validated directly that $[\![w_{14}]\!]^*_{E8} = CX_{0,1}$. Then by application of **Gen**$(+)$, the generator $CX_{0,1}$ is introduced, alongside the relation $CX_{0,1} \approx w_{14}$.

E Establishing the Minimality of $W(E_8)$ and $\mathrm{O}(8,\mathbb{D})$ Generators

This section establishes the minimality of certain generating sets for $W(E_8)$ and $\mathrm{O}(8,\mathbb{D})$. First, a general result about minimal generating sets is established. This result is then applied to the generating sets of interest, to prove their minimality.

E.1 Two Results on Minimal Generating Sets

Theorem E.1. *Let G be a group with $\Sigma' \subseteq \Sigma \subseteq G$. If there exists a $g \in G$ such that g commutes with the elements of Σ' and g does not commute with the elements of Σ, then $\langle \Sigma' \rangle$ is a proper subgroup of $\langle \Sigma \rangle$.*

Proof. Assume that $g \in G$, g commutes with every element of Σ', and $\langle \Sigma' \rangle = \langle \Sigma \rangle$. It follows by induction on the length of an element in $\langle \Sigma' \rangle$, that g commutes with every element in $\langle \Sigma' \rangle$. As a base case, if $h \in \langle \Sigma' \rangle$ corresponds to a word of length 0, then h is the identity and $g \circ h = g = h \circ g$. As an inductive hypothesis, assume that for some $n \in \mathbb{N}$, if $h_1, h_2, \ldots, h_n \in \Sigma'$ and $h = h_1 \circ h_2 \circ \cdots \circ h_n$, then $g \circ h = h \circ g$. To show that the inductive step holds, let $h_1, h_2, \ldots, h_n, h_{n+1} \in \Sigma'$ and $h = h_1 \circ h_2 \circ \cdots h_{n+1}$. By the inductive hypothesis, $g \circ h' = h' \circ g$ where $h' = h_1 \circ h_2 \circ \cdots \circ h_n$. Then $g \circ h = g \circ h' \circ h_{n+1} = h' \circ g \circ h_{n+1} = h' \circ h_{n+1} \circ g = h \circ g$. Then the inductive step holds, and g commutes with every element of $\langle \Sigma' \rangle$. In particular, g commutes with Σ. By the contrapositive, if g does not commute with Σ, then $\langle \Sigma' \rangle \neq \langle \Sigma \rangle$. However, $\langle \Sigma' \rangle \leq \langle \Sigma \rangle$ since $\Sigma' \subseteq \Sigma$. Therefore, $\langle \Sigma' \rangle$ is a proper subgroup of $\langle \Sigma \rangle$. □

Lemma E.2. *Let G be a group with $\Sigma \subseteq G$. If for every maximal proper subset Σ' of Σ, $\langle \Sigma' \rangle$ is a proper subgroup of $\langle \Sigma \rangle$, then Σ is a minimal generating set for $\langle \Sigma \rangle$.*

Proof. Let Σ' be a proper subset of Σ. Then there exists some maximal proper subset Π of Σ such that $\Sigma' \subseteq \Pi \subseteq \Sigma$. Then $\langle \Sigma' \rangle \leq \langle \Pi \rangle \leq \langle \Sigma \rangle$. Since Π is maximal, then by assumption, $\langle \Pi \rangle$ is a proper subgroup of $\langle \Sigma \rangle$. Consequently, $\langle \Sigma' \rangle$ is a proper subgroup of $\langle \Sigma \rangle$. Since Σ' was arbitrary, then Σ is a minimal generating set for $\langle \Sigma \rangle$. □

E.2 Minimality for $W(E_8)$

It must be shown that for every maximal proper subset Σ' of Σ_0, there exists some 8×8 dyadic matrix M such that M commutes with Σ' but does not commute with Σ_0. The first three cases can be solved by inspection. In fact, these matrices follow from well-known circuit relations.

1. Z_2 commutes with $\{X_0, CX_{0,1}, CCX_{1,2}\}$ but does not commute with $K_{1,2}$.
2. H_2 commutes with $\{X_0, CX_{0,1}, K_{1,2}\}$ but does not commute with $CCX_{1,2}$.
3. X_0 commutes with $\{X_0, CCX_{1,2}, K_{1,2}\}$ but does not commute with $CX_{0,1}$.

The final case is less obvious, but can be reduced to solving a linear integer program. Assume that there exists such a matrix M. Since M is dyadic, then there exists some integer matrix N and integer k such that $M = N/2^k$. Clearly, M and N commute with the same matrices. Then M is characterized by the following four equations.

$$X_0 \circ N \neq N \circ X_0 \qquad CX_{0,1} \circ N = N \circ CX_{0,1} \qquad CCX_{1,2} \circ N = N \circ CCX_{1,2} \qquad K_{1,2} \circ N = N \circ K_{1,2}$$

Without loss of generality, $K_{1,2}$ can be replaced by its integral scalar multiple $2 \cdot K_{1,2}$. Then the entries of N can be thought of as 64 integer variables, with each equation yielding 64 linear constraints. Using

Z3 [20] as a solver, the following solution is obtained.

$$N_{0,0} = \begin{bmatrix} 4 & 2 & 2 & 0 \\ 2 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad N_{0,1} = N_{1,0} = N_{1,1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad N = \begin{bmatrix} N_{0,0} & M_{0,1} \\ N_{1,0} & M_{1,1} \end{bmatrix}$$

This establishes Theorem 4.4.

E.3 Minimality of Σ_K for $\mathrm{O}(8,\mathbb{D})$

First, it will be shown that Σ_K is a minimal generating set for $\mathrm{O}(8,\mathbb{D})$. To see that Σ_K generates $\mathrm{O}(8,\mathbb{D})$ simply note that $\Sigma_K \cup \{K_{1,2}\}$ generates $\mathrm{O}(8,\mathbb{D})$ with $K_{1,2} = K_{[0,1,2,3]} \circ X_0 \circ K_{[0,1,2,3]} \circ X_0$. It remains to be shown that for every maximal proper subset Σ' of Σ_K, there exists some 8×8 dyadic matrix M such that M commutes with Σ' but does not commute with Σ_K. The first three cases are also solved by inspection, using well-known circuit relations.

1. Z_2 commutes with $\{X_0, CX_{0,1}, CCX_{1,2}\}$ but does not commute with $K_{[0,1,2,3]}$.
2. H_2 commutes with $\{X_0, CX_{0,1}, K_{[0,1,2,3]}\}$ but does not commute with $CCX_{1,2}$.
3. $X_2 \circ CZ_{0,2} \circ X_2$ commutes with $\{CX_{0,1}, CCX_{1,2}, K_{[0,1,2,3]}\}$ but does not commute with X_0.

Using Z3 as in the $W(E_8)$, it is then possible to find an integer matrix L such that L commutes with $\{X_0, CCX_{1,2}, K_{[0,1,2,3]}\}$ but does not commute with $CX_{0,1}$. The solution is as follows.

$$L_0 = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 2 & 0 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \qquad L_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} L_0 & L_1 \\ L_1 & L_0 \end{bmatrix}$$

This establishes the first claim of Theorem 5.3.

E.4 Minimality of Σ_Z for $\mathrm{O}(8,\mathbb{D})$

The proof the Σ_Z is minimal proves more challenging. However, the minimality of Σ_K can be used to simplify this argument significantly. Clearly, $\Sigma_D = \Sigma_Z \setminus \{CCZ\}$ does not generate $\mathrm{O}(8,\mathbb{D})$, since $\langle \Sigma_D \rangle = W(E_8) < \mathrm{O}(8,\mathbb{D})$. Three of the remaining four cases can be solved by inspection.

1. Recall N from Appendix E.2. By construction, this matrix commutes with $\{CX_{0,1}, CCX_{1,2}, K_{1,2}\}$ and does not commute with X_0. Furthermore, $CCZ \circ N = N \circ CCZ$, since the 8-th row and 8-th column of N contain only zeros.
2. $\sigma_{1,2}$ commutes with $\{X_0, CCX_{1,2}, K_{1,2}, CCZ\}$ but does not commute with $CX_{0,1}$.
3. Z_2 commutes with $\{X_0, CX_{0,1}, CCX_{1,2}, CCZ\}$ but does not commute with $K_{1,2}$.

The case of $CCX_{1,2}$ requires more care. There is no obvious operation which commutes with all generators except for $CCX_{1,2}$. Furthermore, Z3 fails to find an solution to the corresponding integer program. Instead, consider the automorphism $f : M \mapsto H_2 \circ M \circ H_2$ of $\mathrm{O}(8,\mathbb{D})$. Since f fixes $\{X_0, CX_{0,1}, K_{1,2,}\}$ and maps CCZ to $CCX_{0,1}$, then f induces an isomorphism between the subgroups $\langle X_0, CX_{0,1}, K_{1,2}, CCZ \rangle$ and $\langle X_0, CX_{0,1}, K_{1,2}, CCX_{0,1} \rangle$ of $\mathrm{O}(8,\mathbb{D})$. Since $X_0, CX_{0,1}, K_{1,2,}, CCX_{0,1} \in W(E_8)$ with $W(E_8)$ finite, then

$$|\langle X_0, CX_{0,1}, K_{1,2}, CCZ \rangle| = |\langle X_0, CX_{0,1}, K_{1,2}, CCX_{1,2} \rangle| \leq |W(E_8)| < \infty.$$

Since $\mathrm{O}(8,\mathbb{D})$ is an infinite group, then $\langle X_0, CX_{0,1}, K_{1,2}, CCZ \rangle < \mathrm{O}(8,\mathbb{D})$. This establishes the second claim of Theorem 5.3.

F Proof Details for a Presentation of $\mathrm{O}(8,\mathbb{D})$

In Section 5, many informal claims were made about the relations in $\mathscr{G}_n$, and the derivations that are possible using these relations. This section restates each claim as a lemma or theorem, provides a proof for each claim, and then explains how these claims establish the lemmas and theorems in Section 5.

F.1 Counting the Relations in $\mathscr{R}_n$

This section validates the claim that $\mathscr{G}_8$ contains 2039 relations. We say that two relations $(q,r) \in \mathscr{G}_8$ and $(q',r') \in \mathscr{G}_8$ are distinct if $q \neq q'$ or $r \neq r'$. This means, for example, that the relations $K_{[0,1,2,3]} \cdot K_{[4,5,6,7]} \approx K_{[4,5,6,7]} \cdot K_{[4,5,6,7]}$ and $K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot K_{[4,5,6,7]}$ are distinct The techniques used in this section can be generalized to count the number of relations in $\mathscr{G}_n$.

First, consider the relations whose parameters are linearly ordered. If a relation schema r has m linearly ordered parameters, then each choice of m distinct numbers in $[n]$ corresponds to a unique instance of r. It follows that a relation schema with m linearly ordered parameters corresponds to $\binom{n}{m}$ unique relations. For each choice of m, we compute $\binom{8}{m}$ and count the number of relations with m linearly ordered parameters.

- If $m = 1$, then there are $\binom{8}{1} = 8$ instances. The only relation with a single parameter is Relation (48). Then this case contributes 8 relations.
- If $m = 2$, then there are $\binom{8}{2} = 28$ instances. The only relations with two parameters, all linearly ordered, are Relations (47) and (58). Then this case contributes 56 relations.
- If $m = 3$, then there are $\binom{8}{3} = 56$ instances. The only relations with three parameters, all linearly ordered, are Relations (56) and (57). Then this case contributes 102 relations.
- If $m = 4$, then there are $\binom{8}{4} = 70$ instances. The relations Relations (49), (63), (64) and (65) all have exactly four parameters, which are all linearly ordered. Then this case contributes 280 relations.
- If $m = 5$, then there are $\binom{8}{5} = 56$ instances. The relations Relations (59), (60), (61) and (62) all have exactly five parameters, which are all linearly ordered. This this case contributes 224 relations.
- If $m = 6$, then there are $\binom{8}{6} = 28$ instances. The only relation with a six parameters is Relation (66). Then this case contributes 28 relations.
- If $m = 8$, then there are $\binom{8}{8} = 1$ instances. The only relation with eight parameters is Relation (67). Then this case contributes 1 relations.

In total, the relation schemata with linearly ordered parameters contribute 699 instances.

The remaining six schemata induce a partial order on the parameters. For example, in Relation (55), the term $K_{[a,b,c,d]} \cdot K_{[e,f,g,h]}$ indicates that $a < b < c < d$ and $e < f < g < h$. However, the choices of (a,b,c,d) are independent from the choices of (e,f,g,h), except that all choices must be distinct. In this example, there are $\binom{n}{4}$ ways to select the four indices in the first order. Then $n-m$ indices remain, from which there are $\binom{n-4}{4}$ choices. In general, for two independent linear orders with m and k parameters respectively, there will be $\binom{n}{m} \cdot \binom{n-m}{k}$ choices. The six schemata are described below.

- In Relation (54), $m = 1$ and $k = 1$, resulting in $\binom{8}{1} \cdot \binom{7}{1} = 56$ choices.
- In Relation (51), $m = 1$ and $k = 2$, resulting in $\binom{8}{1} \cdot \binom{7}{2} = 168$ choices.
- In Relation (53), $m = 1$ and $k = 4$, resulting in $\binom{8}{1} \cdot \binom{7}{4} = 280$ choices.
- In Relation (50), $m = 2$ and $k = 2$, resulting in $\binom{8}{2} \cdot \binom{6}{2} = 420$ choices.

– In Relation (52), $m = 2$ and $k = 4$, resulting in $\binom{8}{2} \cdot \binom{6}{4} = 420$ choices.

– In Relation (55), $m = 4$ and $k = 4$, resulting in $\binom{8}{4} \cdot \binom{4}{4} = 70$ choices

In total, the relations partially ordered parameters contribute 1414 instances. Then $|\mathscr{R}_8| = 2113$.

F.2 Correctness of Relation Reindexing

This section justifies the reindexing of relations via permutations. First, recall that every permutation on $[n]$ can be represented by a permutation of the basis vectors in $\mathbb{R}^8$, with $\tau_{a,b}$ corresponding to $X_{[a,b]}$. The intuition is that every σ can be represented by a word w over generators of type X, and that conjugation by w corresponds to formal reindexing when σ is valid.

First, a subset $\mathscr{R}_n^B$ of $\mathscr{R}_n$ is identified, for which all of the order and braiding relations for generators of type X hold. Consequently, $\mathscr{R}_n^B$ is complete for words over generators of type X. Then $\mathscr{R}_n^B$ is extended to a subset $\mathscr{R}_n^\sigma$ of $\mathscr{R}_n$ for which all formal reindexings are derivable. The result is proven first, for individual generators, and then extended to entire words.

F.2.1 Deriving the Braiding Relations

First, define the set of relations,

$$\mathscr{R}_n^\tau = \left\{ {X_{[a,a+1]}}^2 \approx \varepsilon \mid a \in \mathbb{Z} \right\} \cup \left\{ X_{[a,a+1]} \cdot X_{[a,b]} \approx X_{[a+1,b]} \cdot X_{[a,a+1]} \mid a,b \in \mathbb{Z} \text{ with } a+1 < b \right\}.$$

This set is sufficient to decompose all swaps into transpositions, as proven in Lemma F.1. Of interest in this section is the following extension of $\mathscr{R}_n^\tau$,

$$\mathscr{R}_n^B = \mathscr{R}_n^\tau \cup \left\{ X_{[a+1,a+2]} \cdot X_{[a,a+1]} \approx X_{[a,a+2]} \cdot X_{[a+1,a+2]} \mid a \in \mathbb{Z} \right\}.$$

The relations in $\mathscr{R}_n^B$ entail the braiding and order relations for $S(n)$, as shown in Lemma F.2, and are therefore complete for equality of words over generators of type X. Of important note is that $\mathscr{R}_n^B \subseteq \mathscr{R}_n$.

Lemma F.1. *Let R be a set of relations over $\mathscr{G}_n$ which contains all well-formed relations in $\mathscr{R}_n^\tau$. If v is a word over two-level operators of type X, then there exists a word u over transpositions such that $v \sim_R u$.*

Proof. Consider a valid two-level operator $X_{[a,b]}$. The proof follows by induction on $b - a > 0$.

– **Base Case**. If $b - a = 1$, then $X_{[a,b]} = X_{[a,a+1]}$.

– **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}_{>0}$, if $b - a = k$, then there exists a word w

– **Inductive Step**. Assume that $b - a = k + 1$. Since $k > 0$, then $a + 1 \neq b$, and the following derivation holds.

$$X_{[a,b]} \leftarrow {X_{[a,a+1]}}^2 \cdot X_{[a,b]} \rightarrow X_{[a,a+1]} \cdot X_{[a+1,b]} \cdot X_{[a,a+1]}$$

Since $b - (a+1) = k$, then by the inductive hypothesis, there exists a word w over transpositions such that $X_{[a+1,b]} \sim_R w$. Then $X_{[a,b]} \sim_R X_{[a,a+1]} \cdot w \cdot X_{[a,a+1]}$. Since $X_{[a,a+1]} \cdot w \cdot X_{[a,a+1]}$ is a word over transpositions, then the inductive hypothesis holds.

Then for each symbol M in w, there exists a decomposition of M into transpositions. Then by Appendix B.2, there exists a word v over transpositions such that $w \sim_R v$. □

Lemma F.2. *Let R be a set of relations over $\mathscr{G}_n$ which contains all well-formed relations in $\mathscr{R}_n^B$. If v and w are words over two-level operators of type X and $[\![v]\!]_O^* = [\![w]\!]_O^*$, then $v \sim_{\mathscr{R}_\sigma} w$.*

Proof. Since v and w are words over two-level operators of type X with $\mathscr{R}_n^\tau \subseteq \mathscr{R}_n^B$, then by Lemma F.1 there exists words $\hat{v}$ and $\hat{w}$ over transpositions such that $\hat{v} \sim \hat{w}$. Then $\hat{v}$ and $\hat{w}$ are words in the braid representation of $\mathrm{S}(n)$. If R contains all order and braiding relations for the transpositions in $\mathscr{G}_n$, then R is complete for words over the transpositions in $\mathscr{G}_n$. Let $a \in [n-2]$. Then the following derivation holds using the relations in $\mathscr{R}_n^B$.

$$X_{[a,a+1]} \cdot X_{[a+1,a+2]} \cdot X_{[a,a+1]} \leftarrow {X_{[a,a+1]}}^2 \cdot X_{[a,a+2]} \rightarrow X_{[a,a+2]} \leftarrow X_{[a,a+2]} \cdot {X_{[a+1,a+2]}}^2 \leftarrow X_{[a+1,a+2]} \cdot X_{[a,a+1]} \cdot X_{[a+1,a+2]}$$

Then $X_{[a,a+1]} \cdot X_{[a+1,a+2]} \cdot X_{[a,a+1]} \sim_R X_{[a+1,a+2]} \cdot X_{[a,a+1]} \cdot X_{[a+1,a+2]}$. Since a was arbitrary, then R is complete for $\mathrm{S}(n)$. Since $v \sim_R \hat{v}$ and $w \sim_R \hat{w}$, then $[\![v]\!]_O^* = [\![\hat{v}]\!]_O^*$ and $[\![w]\!]_O^* = [\![\hat{w}]\!]_O^*$. Then $[\![\hat{v}]\!]_O^* = [\![\hat{w}]\!]_O^*$. Then $\hat{v} \sim_R \hat{w}$ by the completeness of R. Then $v \sim_R w$ by the transitivity of $(\sim_R)$. □

F.2.2 Formal Inverses for Self-Inverse Generators

For each $w = w_1 \cdot w_2 \cdots w_n$ over $\mathscr{G}_n$, define $\overline{w} = w_n \cdots w_2 \cdot w_1$. Since each element in $\mathscr{G}_n$ is self-inverse, then $[\![\overline{w}]\!]_O^*$ is the inverse of $[\![w]\!]_O^*$ in $\mathrm{O}(8,\mathbb{D})$. One can prove that given a complete set of relations, both $u \cdot \overline{u}$ and $\overline{u} \cdot u$ always derive to ε. For the purposes of this proof, only the case for X-type generators is necessary.

Lemma F.3. *Let R be a set of relations over $\mathscr{G}_n$ which contains all well-formed relations in the set below.*

$$\{{X_{[a,a+1]}}^2 \approx \varepsilon \mid a \in \mathbb{Z}\} \cup \{X_{[a,a+1]} \cdot X_{[a,b]} \approx X_{[a+1,b]} \cdot X_{[a,a+1]} \mid a,b \in \mathbb{Z}\} \cup \{X_{[a+1,a+2]} \cdot X_{[a,a+1]} \approx X_{[a,a+@]} \cdot X_{[a+1,a+2]} \mid a \in \mathbb{Z}\}$$

If u is a word over two-level operators of type X, then $u \cdot \overline{u} \sim_R \varepsilon$ and $\overline{u} \cdot u \sim_R \varepsilon$. Furthermore, if v is a word over two-level operators of type X and $u \sim_R v$, then $\overline{u} \sim_R \overline{v}$.

Proof. Since $[\![\cdot]\!]_O^*$ maps each generator in $\mathscr{G}_n$ to a self-inverse matrix, then $[\![\overline{u}]\!]_O^*$ is the inverse to $[\![u]\!]_O^*$. Then $[\![\overline{u} \cdot u]\!]_O^* = [\![\varepsilon]\!]_O^* = [\![u \cdot \overline{u}]\!]_O^*$. Since u is a word over two-level operators of type X, then $u \cdot \overline{u} \sim_R \varepsilon$ and $\overline{u} \cdot u \sim_R \varepsilon$ by Lemma F.2. Now assume that v is a word over two-level operators of type X with $u \sim_R v$. Then $[\![u]\!]_O^* = [\![v]\!]_O^*$. Since $\overline{u}$ is the inverse to u and $\overline{v}$ is the inverse to v, then $[\![\overline{u}]\!]_O^* = [\![\overline{v}]\!]_O^*$. Then $\overline{u} \sim_R \overline{v}$ by Lemma F.2. □

F.2.3 Permuting the Indices in Multi-Level Operators

Theorem F.4. *If $\sigma \in \mathrm{S}(n)$ is a valid reindexing for a two-level operator M of type X, then there exists a word v over the two-level operators of type X, such that $[\![v]\!]_O^* = [\![\sigma]\!]_S$ and $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since M is a two-level operator of type X, then there exists an increasing sequence (a,b) over $[n]$ such that $M = X_{[a,b]}$. Let $\sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$ be the decomposition of σ into a sequence of transpositions. Then define $v = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_m]\!]_S$. Clearly v is a word over two-level operators of type X satisfying $[\![v]\!]_O^* = [\![\sigma]\!]_S$. Furthermore, $[\![\sigma(M)]\!]_O^* = [\![\tau_{\sigma(a),\sigma(b)}]\!]_S = [\![\sigma \cdot \tau_{a,b} \cdot \sigma^{-1}]\!]_S^* = [\![v \cdot M \cdot \overline{v}]\!]_O^*$. Since $\sigma(M)$ and $\overline{v} \cdot M \cdot v$ are words over two-level operators of type X, then $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ by Lemma F.2. □

Theorem F.5. *If $\sigma \in \mathrm{S}(n)$ and M is a one-level operator of type (-1), then there exists a word v of transpositions, such that $[\![v]\!]_O^* = [\![\sigma]\!]_S$ and $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since M is a one-level operator of type (-1), then there exists an $a \in [a]$ such that $M = (-1)_{[a]}$. Let v Let $\sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$ be the decomposition of σ into a sequence of transpositions. Then define $v = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_m]\!]_S$. It follows by induction on m that $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.

– **Base Case**. Assume that $m = 0$. Then $v = \overline{v} = \varepsilon$ and $\sigma(M) = M$. Then $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ by the reflexivity of $(\sim_{\mathscr{R}_\sigma})$.

– **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $m = k$, then $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.

– **Inductive Step**. Assume that $m = k+1$ and define $u = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_k]\!]_S$. Then by the inductive hypothesis $\sigma(M) \sim_{\mathscr{R}_\sigma} u \cdot \sigma_m(M) \cdot \overline{u}$. Since σ_m is a transposition, then there exists some $j \in [n-1]$ such that $\sigma_m = \tau_{j,j+1}$. Furthermore, $[\![\sigma_m]\!]_S = X_{[j,j+1]}$. If $j = a$, then $\sigma_m(M) = (-1)_{[a+1]}$ then the following derivation holds using only Relations (47) and (58).

$$(-1)_{[a+1]} \leftarrow (-1)_{[a+1]} \cdot {X_{[j,j+1]}}^2 \leftarrow X_{[j,j+1]} \cdot (-1)_{[a]} \cdot X_{[j+1]}$$

The case when $j+1 = a$ follows symmetrically. When $j \neq a$ and $j+1 \neq a$, then $\sigma_m(M) = M$ and the following derivation holds using only Relations (47) and (51).

$$(-1)_{[a]} \leftarrow (-1)_{[a]} \cdot {X_{[j,j+1]}}^2 \leftarrow X_{[j,j+1]} \cdot (-1)_{[a]} \cdot X_{[j+1]}$$

In either case, $\sigma_m(M) \sim_{\mathscr{R}_\sigma} [\![\sigma_m]\!]_S \cdot M \cdot [\![\sigma_m]\!]_S$. Then $\sigma_m(M) \sim_{\mathscr{R}_\sigma} u \cdot \sigma_m(M) \cdot \overline{u} \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ and the inductive step is established.

Then by the principle of induction, $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$. Clearly $[\![v]\!]^*_O = [\![\sigma]\!]_S$. □

Lemma F.6. *For each four-level operator M of type K, there exists a valid reindexing σ for M and a word v over two-level operators of type X, such that $[\![v]\!]^*_O = [\![\sigma]\!]_S$ and $K_{[0,1,2,3]} \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since M is a four-level operator of type K, then there exists an increasing sequence (a_0, a_1, a_2, a_3) over $[n]$ such that $M = (-1)_{[a_0,a_1,a_2,a_3]}$. Since (a_0, a_1, a_3, a_4) is increasing, then $k \leq a_k$ for $k \in [4]$. Then for each $k \in [4]$, define σ_k to be τ_{k,a_k} if $k \neq a_k$, or identity otherwise. Then the following equations hold.

$$\sigma_0(M) = (-1)_{[0,a_1,a_2,a_3]} \qquad \sigma_1(\sigma_0(M)) = (-1)_{[0,1,a_2,a_3]}$$
$$\sigma_2(\sigma_1(\sigma_0(M))) = (-1)_{[0,1,2,a_3]} \qquad \sigma_3(\sigma_2(\sigma_1(\sigma_0(M)))) = (-1)_{[0,1,2,3]}$$

Let $v = [\![\sigma_3]\!]_S \cdot [\![\sigma_2]\!]_S \cdot [\![\sigma_1]\!]_S \cdot [\![\sigma_0]\!]_S$. Then the following derivations hold by Relations (47), (59), (60), (61) and (62). We assume that each σ_k is not the identity, else the derivation is trivial.

$$\sigma_0(M) \leftarrow {X_{[0,a_0]}}^2 \cdot K_{[0,a_1,a_2,a_3]} \rightarrow X_{[0,a_0]} \cdot K_{[a_0,a_1,a_2,a_3]} \cdot X_{[0,a_0]}$$
$$\sigma_1(\sigma_0(M)) \leftarrow {X_{[1,a_1]}}^2 \cdot K_{[0,1,a_2,a_3]} \rightarrow X_{[1,a_1]} \cdot K_{[0,a_1,a_2,a_3]} \cdot X_{[1,a_1]} = X_{[1,a_1]} \cdot \sigma_0(M) \cdot X_{[1,a_1]}$$
$$\sigma_2(\sigma_1(\sigma_0(M))) \leftarrow {X_{[2,a_2]}}^2 \cdot K_{[0,1,2,a_3]} \rightarrow X_{[2,a_2]} \cdot K_{[0,1,a_2,a_3]} \cdot X_{[2,a_2]} = X_{[2,a_2]} \cdot \sigma_1(\sigma_0(M)) \cdot X_{[2,a_2]}$$
$$\sigma_3(\sigma_2(\sigma_1(\sigma_0(M)))) \leftarrow {X_{[3,a_3]}}^2 \cdot K_{[0,1,2,3]} \rightarrow X_{[3,a_3]} \cdot K_{[0,1,2,a_3]} \cdot X_{[3,a_3]} = X_{[3,a_3]} \cdot \sigma_1(\sigma_0(M)) \cdot X_{[3,a_3]}$$

It follows that $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ where $\sigma = \sigma_3 \cdot \sigma_2 \cdot \sigma_1 \cdot \sigma_0$. Clearly $[\![v]\!]^*_O = [\![\sigma]\!]_S$. □

Lemma F.7. *Let $M = K_{[0,1,2,3]}$ be a four-level operator of dimension n. For any increasing sequence (a_0, a_1, a_2, a_3) over $[n]$, there exists a valid reindexing σ for M and a word v over two-level operators of type X, such that $\sigma(M) = K_{[a_0,a_1,a_2,a_3]}$, $[\![v]\!]^*_O = [\![\sigma]\!]_S$, and $K_{[a_0,a_1,a_2,a_3]} \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since (a_0, a_1, a_3, a_4) is increasing, then $k \leq a_k$ for $k \in [4]$. Then for each $k \in [4]$, define σ_k to be τ_{k,a_k} if $k \neq a_k$, or identity otherwise. Then the following equations hold.

$$\sigma_3(M) = (-1)_{[0,1,2,a_3]} \qquad \sigma_2(\sigma_3(M)) = (-1)_{[0,1,a_2,a_3]}$$
$$\sigma_1(\sigma_2(\sigma_3(M))) = (-1)_{[0,a_1,a_2,a_3]} \qquad \sigma_0(\sigma_1(\sigma_2(\sigma_3(M)))) = (-1)_{[a_0,a_2,a_3,a_4]}$$

Let $v = [\![\sigma_0]\!]_S \cdot [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdot [\![\sigma_3]\!]_S$. Then the following derivations hold by Relations (47), (59), (60), (61) and (62). We assume that each σ_k is not the identity, else the derivation is trivial.

$$\sigma_3(M) \leftarrow K_{[0,1,2,3]} \cdot X_{[3,a_3]}{}^2 \leftarrow X_{[3,a_3]} \cdot K_{[0,1,2,a_3]} \cdot X_{[3,a_3]}$$
$$\sigma_2(\sigma_3(M)) \leftarrow K_{[0,1,a_2,a_3]} \cdot X_{[2,a_2]}{}^2 \leftarrow X_{[2,a_2]} \cdot K_{[0,1,2,a_3]} \cdot X_{[2,a_2]} = X_{[2,a_2]} \cdot \sigma_3(M) \cdot X_{[2,a_2]}$$
$$\sigma_1(\sigma_2(\sigma_3(M))) \leftarrow K_{[0,a_1,a_2,a_3]} \cdot X_{[1,a_1]}{}^2 \leftarrow X_{[1,a_1]} \cdot K_{[0,1,a_2,a_3]} \cdot X_{[1,a_1]} = X_{[1,a_1]} \cdot \sigma_2(\sigma_3(M)) \cdot X_{[1,a_1]}$$
$$\sigma_0(\sigma_1(\sigma_2(\sigma_3(M)))) \leftarrow K_{[a_0,a_1,a_2,a_3]} \cdot X_{[0,a_0]}{}^2 \leftarrow X_{[0,a_0]} \cdot K_{[0,a_1,a_2,a_3]} \cdot X_{[0,a_0]} = X_{[0,a_0]} \cdot \sigma_1(\sigma_2(\sigma_3(M))) \cdot X_{[0,a_0]}$$

It follows that $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ where $\sigma = \sigma_0 \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3$. Clearly $[\![v]\!]^*_O = [\![\sigma]\!]_S$. □

Lemma F.8. *If $\sigma \in \mathrm{S}(n)$ is a valid reindexing for a four-level operator M of type K and $\sigma(M) = M$, then there exists a word v over the two-level operators of type X, such that $[\![v]\!]^*_O = [\![\sigma]\!]_S$ and $M \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since M is a four-level operator of type K, then there exists an increasing sequence (a,b,c,d) over $[n]$ such that $M = (-1)_{[a,b,c,d]}$. Since $\sigma(M) = M$, then σ fixes $\{a,b,c,d\}$. Then σ restricts to a permutation on $[n] \setminus \{a,b,c,d\}$. Decompose this restriction of σ into a sequence of transpositions $\sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$ on $[n] \setminus \{a,b,c,d\}$. Since σ fixes $\{a,b,c,d\}$, then $\sigma = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$ when viewing each σ_j as a permutation on $[n]$. Define $v = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_m]\!]_S$. It follows by induction on m that $M \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.

- **Base Case**. If $m = 0$, then $v = \overline{v} = \varepsilon$. Then $M \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$ by the reflexivity of $(\sim_{\mathscr{R}_\sigma})$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $m = k$, then $M \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.
- **Inductive Step**. Assume that $m = k+1$ and define $u = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_k]\!]_S$. Then by the inductive hypothesis $M \sim_{\mathscr{R}_\sigma} u \cdot M \cdot \overline{u}$. Since σ_m is a transposition of elements in $[n] \setminus \{a,b,c,d\}$, then there exists some $j,l \in [n] \setminus \{a,b,c,d\}$ such that $\sigma_m = \tau_{j,l}$. Furthermore, $[\![\sigma_m]\!]_S = X_{[j,l]}$. Since $j,l \notin \{a,b,c,d\}$, then the following derivation holds by Relations (47) and (52).

  $$K_{[a,b,c,d]} \leftarrow X^2_{[j,k]} \cdot K_{[a,b,c,d]} \rightarrow X_{[j,l]} \cdot K_{[a,b,c,d]} \cdot X_{[j,l]}$$

  Then $M \sim_{\mathscr{R}_\sigma} [\![\sigma_m]\!]_S \cdot M \cdot [\![\sigma_m]\!]_S$. Then $M \sim_{\mathscr{R}_\sigma} u \cdot M \cdot \overline{u} \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.

Then by the principle of induction, $M \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$. Clearly $[\![v]\!]^*_O = [\![\sigma]\!]_S$. □

Theorem F.9. *If $\sigma \in \mathrm{S}(n)$ is a valid reindexing for a four-level operator M of type K, then there exists a word v over the two-level operators of type X, such that $[\![v]\!]^*_O = [\![\sigma]\!]_S$ and $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$.*

Proof. Since M is a four-level operator of type K, then there exists an increasing sequence (a_0,a_1,a_2,a_3) over $[n]$ such that $M = (-1)_{[a_0,a_1,a_2,a_3]}$. By Lemma F.6, there exists a word u over two-level operators of type X, and a permutation σ_1 such that $\sigma_1(M) = (-1)_{[0,1,2,3]}$, $[\![\sigma_1]\!]_S = [\![u]\!]^*_S$, and $\sigma_1(M) \sim_{\mathscr{R}_\sigma} u \cdot M \cdot \overline{u}$. By Lemma F.7, there exists a word v over two-level operators of type X, and a permutation σ_2 such that $\sigma_2(\sigma_1(M)) = (-1)_{[\sigma(a),\sigma(b),\sigma(b),\sigma(c)]} = \sigma(M)$, $[\![\sigma_2]\!]_S = [\![v]\!]^*_S$, and $\sigma_2(\sigma_1(M)) \sim_{\mathscr{R}_\sigma} v \cdot \sigma_1(M) \cdot \overline{v}$. Then define $\sigma_3 = \sigma \circ \sigma_1^{-1} \circ \sigma_2^{-1}$. Then $\sigma_3(\sigma(a_k)) = \sigma(\sigma_1^{-1}(\sigma_2^{-1}(\sigma(a_k)))) = \sigma(\sigma_1^{-1}(k)) = \sigma(a_k)$ for all $k \in [4]$. Then σ_3 is a valid reindexing for M with $\sigma_3(\sigma_2(\sigma_1(M))) = M$. By Lemma F.8, there exists a word w over two-level operators of type X, such that $[\![\sigma_3]\!]_S = [\![w]\!]^*_S$ and $\sigma(M) \sim_{\mathscr{R}_\sigma} w \cdot \sigma_2(\sigma_1(M)) \cdot \overline{w}$. It follows that $\sigma(M) \sim_{\mathscr{R}_\sigma} w \cdot \sigma_2(\sigma_1(M))\overline{w} \sim_{\mathscr{R}_\sigma} w \cdot u \cdot \sigma_1(M) \cdot \overline{w \cdot u} \sim_{\mathscr{R}_\sigma} w \cdot u \cdot v \cdot M \cdot \overline{w \cdot u \cdot v}$. Moreover, $[\![w \cdot u \cdot v]\!]^*_O = [\![w]\!]^*_O \circ [\![u]\!]^*_O \circ [\![v]\!] = [\![\sigma_3]\!]_S \circ [\![\sigma_2]\!]_S \circ [\![\sigma_1]\!]_S = [\![\sigma]\!]_S$. □

F.2.4 Permuting the Indices in Relations Over Multi-Level Operators

Theorem F.10. *For each $\sigma \in \mathrm{S}(n)$, there exits a set of words L_σ with the following properties.*

1. *If $v_1 \in L_\sigma$ and $v_2 \in L_\sigma$, then $v_1 \sim_{\mathscr{R}_\sigma} v_2$.*
2. *If σ is a valid reindexing for w, then there exists a $v \in L_\sigma$ such that $\sigma(w) \sim_{\mathscr{G}_\sigma} v \cdot w \cdot \overline{v}$.*

Proof. Let L_σ be the set of all words v over the two-level operators of type X, such that $[\![v]\!]^*_O = [\![\sigma]\!]_S$. Let $v_1 \in L_\sigma$ and $v_2 \in L_\sigma$. Then $[\![v_1]\!]^*_O = [\![\sigma]\!]_S = [\![v_2]\!]^*_O$. Then $v_1 \sim_{\mathscr{G}_\sigma} v_2$. Since v_1 and v_2 were arbitrary, then Property (1) holds. Now assume that σ is a valid reindexing for w. Property (2) follows by induction on the length of w.

- **Base Case**. If $|w| = 0$, then $\sigma(w) = w$. Let $\sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$ be a decomposition of σ into transpositions. Define $v = [\![\sigma_1]\!]_S \cdot [\![\sigma_2]\!]_S \cdots [\![\sigma_m]\!]$. Then $[\![\overline{v}]\!]^*_O = [\![\sigma^{-1}]\!]_S$ since $\overline{v}$ is the inverse to v. Since $v \cdot w \cdot \overline{v}$, then $\sigma(w) \sim_{\mathscr{R}_\sigma} v \cdot w \cdot \overline{v}$ by Lemma F.3.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|w| = k$, then there exists a $v \in L_\sigma$ such that $\sigma(w) \sim_{\mathscr{G}_\sigma} v \cdot w \cdot \overline{v}$.
- **Inductive Step**. Assume that $|w| = k+1$. Then there exists some word $\hat{w}$ over $\mathscr{G}_n$ and some $M \in \mathscr{G}_n$ such that $w = \hat{w} \cdot M$ with $|\hat{w}| = k$. Clearly, M is either of type X, type (-1), or type K. In any case, there exists a word v over the two-level operators of type X such that $[\![v]\!]^*_O = [\![\sigma]\!]_S$ and $\sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot M \cdot \overline{v}$. Then $v \in L_\sigma$. By the inductive hypothesis, there exists a $u \in L_\sigma$ such that $\sigma(\hat{w}) \sim_{\mathscr{R}_\sigma} u \cdot \hat{w} \cdot \overline{u}$. Since $u \in L_\sigma$ and $v \in L_\sigma$, then $u \sim_{\mathscr{R}_\sigma} v$ by Property (1). Furthermore $\overline{v} \cdot v \sim_{\mathscr{R}_\sigma} \varepsilon$ by Lemma F.3. Then $\overline{v} \cdot u \sim_{\mathscr{R}_\sigma} \overline{v} \cdot v \sim_{\mathscr{R}_\sigma} \varepsilon$. Since $\sigma(w) = \sigma(\hat{w}) \cdot \sigma(M)$, then it follows $\sigma(w) \sim_{\mathscr{R}_\sigma} v \cdot \hat{w} \cdot \overline{v} \cdot \sigma(M) \sim_{\mathscr{R}_\sigma} v \cdot \hat{w} \cdot \overline{v} \cdot u \cdot M \cdot \overline{u} \sim_{\mathscr{R}_\sigma} v \cdot \hat{w} \cdot M \cdot \overline{u} \sim_{\mathscr{R}_\sigma} u \cdot w \cdot \overline{u}$ and the inductive step is established.

Then by the principle of induction, Property (2) holds. □

Corollary F.11. *Let v and w be words over $\mathscr{G}_n$. If σ is a valid reindexing for u and w, then $\sigma(w)$ is derivable from $\sigma(u)$ using $\mathscr{R}_\sigma \cup \{u \approx w\}$.*

Proof. Let $Q = \mathscr{R}_\sigma \cup \{u \approx w\}$. By Theorem F.10, there exists words v_1 and v_2 over two-level operators of type X, such that $v_1 \sim_Q v_2$, $\sigma(u) \sim_Q v_1 \cdot u \cdot \overline{v}_1$ and $\sigma(w) \sim_Q v_2 \cdot w \cdot \overline{v}_2$. Since $v_1 \sim_Q v_2$, then $[\![v_1]\!]^*_O = [\![v_2]\!]^*_O$. Then $[\![\overline{v}_1]\!]^*_O = [\![\overline{v}_2]\!]^*_O$. Since $\overline{v}_1$ and $\overline{v}_2$ are words over two-level operators of type X, then $\overline{v}_1 \sim_Q \overline{v}_2$ by Lemma F.2. Then the following derivation holds over Q.

$$\sigma(u) \to v_1 \cdot u \cdot \overline{v}_1 \to v_2 \cdot u \cdot \overline{v}_1 \to v_2 \cdot w \cdot \overline{v}_1 \to v_2 \cdot w \cdot \overline{v}_2 \to \sigma(w)$$

Then $\sigma(w)$ is derivable from $\sigma(w)$ using Q. □

F.3 The Set of Representative Relations

In Section 5.2, a set of representative relations were selected from $\mathscr{R}_n$. These relations are illustrated in Figure 10. In some sense, the choice of representative relations were arbitrary, since all choices are equivalent up to permutation. However, preference was given to the parameters $[0]$, $[4]$, $[0,1,2,3]$, $[4,5,6,7]$, since these correspond well to controlled qubit operators.

$$X_{[a,b]}{}^2 \approx \varepsilon \quad (109)$$
$$(-1)_{[0]}{}^2 \approx \varepsilon \quad (110)$$
$$K_{[0,1,2,3]}{}^2 \approx \varepsilon \quad (111)$$
$$X_{[a,b]} \cdot X_{[c,d]} \approx X_{[c,d]} \cdot X_{[a,b]} \quad (112)$$
$$X_{[a,b]} \cdot (-1)_{[c]} \approx (-1)_{[c]} \cdot X_{[a,b]} \quad (113)$$
$$X_{[a,b]} \cdot K_{[c,d,e,f]} \approx K_{[c,d,e,f]} \cdot X_{[a,b]} \quad (114)$$
$$(-1)_{[4]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot (-1)_{[4]} \quad (115)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \approx (-1)_{[4]} \cdot (-1)_{[0]} \quad (116)$$
$$K_{[0,1,2,3]} \cdot K_{[4,5,6,7]} \approx K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \quad (117)$$
$$X_{[a,a+2]} \cdot X_{[a,a+1]} \approx X_{[a+1,a+2]} \cdot X_{[a,a+2]} \quad (118)$$
$$X_{[a+1,a+2]} \cdot X_{[a,a+1]} \approx X_{[a,a+2]} \cdot X_{[a+1,a+2]} \quad (119)$$
$$X_{[a,b]} \cdot (-1)_{[a]} \approx (-1)_{[b]} \cdot X_{[a,b]} \quad (120)$$
$$X_{[a,e]} \cdot K_{[a,b,c,d]} \approx K_{[e,b,c,d]} \cdot X_{[a,e]} \quad (121)$$
$$X_{[b,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,e,c,d]} \cdot X_{[b,e]} \quad (122)$$
$$X_{[c,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,e,d]} \cdot X_{[c,e]} \quad (123)$$
$$X_{[d,e]} \cdot K_{[a,b,c,d]} \approx K_{[a,b,c,e]} \cdot X_{[d,e]} \quad (124)$$
$$X_{[0,1]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[0,1]} \cdot (-1)_{[1]} \cdot (-1)_{[3]} \quad (125)$$
$$X_{[1,2]} \cdot K_{[0,1,2,3]} \approx (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \cdot K_{[0,1,2,3]} \cdot (-1)_{[0]} \quad (126)$$
$$X_{[2,3]} \cdot K_{[0,1,2,3]} \approx K_{[0,1,2,3]} \cdot X_{[1,3]} \quad (127)$$
$$K_{[0,1,2,3]} \cdot K_{[1,3,4,5]} \approx K_{[1,3,4,5]} \cdot K_{[0,1,2,3]} \quad (128)$$
$$(-1)_{[0]} \cdot (-1)_{[4]} \cdot X_{[0,4]} \cdot \rho \approx \rho \cdot X_{[0,4]} \cdot (-1)_{[4]} \cdot (-1)_{[0]} \quad (129)$$

Figure 10: The representative relations in $\mathscr{R}_n^1$, for all valid choices of $a,b,c,d,e,f \in \mathbb{Z}$. We write ρ for the substring $K_{[4,5,6,7]} \cdot K_{[0,1,2,3]} \cdot X_{[3,4]} \cdot K_{[0,1,2,3]} \cdot K_{[4,5,6,7]} \cdot X_{[0,4]}$.

F.4 Proving the Redundant Relations are Derivable

This section makes use of the braiding relations and the inverse relations, to derive several bifunctoriality and commutator relations. Each proof follows the same structure. First, the special case is proven where all generators of type (-1) or K have consecutive indices starting from 0. In all other cases, there is a generator of type (-1) or K conjugated by a permutation. The braiding relations are used to obtain a convenient decomposition for each permutation. The commutativity and bifunctoriality follow immediately from these decompositions.

Lemma F.12. *If $(0,a,b)$ is an increasing sequence over $[n]$, then $X_{[a,b]} \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot X_{[a,b]}$.*

Proof. Let $\sigma = \tau_{a,b}$. Since $3 < a < b$, then there exists a decomposition $\tau_{c_1,c_1+1} \circ \tau_{c_2,c_2+1} \circ \cdots \tau_{c_m,c_m+1}$ of σ into transpositions such that $0 < c_k$ for all $k \in [n]$. Define $u = X_{[c_1,c_1+1]} \cdot X_{[c_2,c_2+1]} \cdots X_{[c_m,c_m+1]}$. Then $[\![X_{[a,b]}]\!]_O^* = [\![\sigma]\!]_S = [\![u]\!]_O^*$. Then $X_{[a,b]} \sim_{\mathscr{R}_n^3} u$ by Lemma F.2. The proof follows by induction on m.

- **Base Case**. If $|u| = 0$, then $u \cdot (-1)_{[0]} = (-1)_{[0]} \cdot u$. Then $u \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot u$ by reflexivity.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|u| = k$, then $u \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot u$.
- **Inductive Step**. Assume that $m = k+1$. Define $v = X_{[c_1,c_1+1]} \cdot X_{[c_2,c_2+1]} \cdots X_{[c_k,c_k+1]}$. Then by definition $u = v \cdot X_{[c_m,c_m+1]}$. Since $c_m > 0$, then $u \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} v \cdot (-1)_{[0]} \cdot X_{[c_m,c_m+1]}$ by the relation $X_{[c_m,c_m+1]} \cdot (-1)_{[0]} \approx_{\mathscr{R}_n^3} (-1)_{[0]} \cdot X_{[c_m,c_m+1]}$. Furthermore, since $|v| = k$, then by the inductive hypothesis $v \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot v$. Then $v \cdot (-1)_{[0]} \cdot X_{[c_m,c_m+1]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot u$. Then by the transitivity of $(\sim_{\mathscr{R}_n^3})$, $u \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot u$ and the inductive step is established.

Then $u \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot u$ by the principle of induction. Since $X_{[a,b]} \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} u \cdot (-1)_{[0]}$ and $(-1)_{[0]} \cdot u \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot X_{[a,b]}$, then $X_{[a,b]} \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot X_{[a,b]}$ by the transitivity of $(\sim_{\mathscr{R}_n^3})$. □

Theorem F.13. *All instances of Relation (51) are derivable from $\mathscr{R}_n^3$.*

Proof. Let $\{a,b,c\} \in [n]$. Define $\sigma \in \mathrm{S}(n)$ such that σ is τ_{k,c_k} if $c > 0$, or identity otherwise. Likewise, define u to be $X_{[0,c]}$ if $c > 0$, or ε otherwise. Clearly $\tau_{a,b} \circ \sigma = \sigma \circ \tau_{\sigma(a),\sigma(b)}$. Then,

$$[\![X_{[a,b]} \cdot u]\!]_O^* = [\![\tau_{a,b} \circ \sigma]\!]_S = [\![\sigma \circ \tau_{\sigma(a),\sigma(b)}]\!]_S = [\![u \cdot X_{[\sigma(a),\sigma(b)]}]\!]_O^*.$$

Then $X_{[a,b]} \cdot u \sim_{\mathscr{R}_n^3} u \cdot X_{[\sigma(a),\sigma(b)]}$ by Lemma F.2. Likewise, $X_{[\sigma(a),\sigma(b)]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} \overline{u} \cdot X_{[a,b]}$ by Lemma F.3. Since $\{a,b,c\}$ are distinct, then $\sigma(a) > 0$ and $\sigma(b) > 0$. Then $X_{[\sigma a,\sigma b]} \cdot (-1)_{[0]} \sim_{\mathscr{R}_n^3} (-1)_{[0]} \cdot X_{[\sigma a,\sigma b]}$ by Lemma F.12. Then the following derivation holds.

$$X_{[a,b]} \cdot u \cdot (-1)_{[0]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot X_{[\sigma(a),\sigma(b)]} \cdot (-1)_{[0]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot (-1)_{[0]} \cdot X_{[\sigma(a),\sigma(b)]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot (-1)_{[0]} \cdot \overline{u} \cdot X_{[a,b]}$$

Since $\{a,b,c\}$ were arbitrary, then all instances of Relation (51) are derivable from $\mathscr{R}_n^3$. □

Lemma F.14. *If $(3,a,b)$ is an increasing sequence over $[n]$, then $X_{[a,b]} \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[a,b]}$.*

Proof. Let $\sigma = \tau_{a,b}$. Since $3 < a < b$, then there exists a decomposition $\tau_{c_1,c_1+1} \circ \tau_{c_2,c_2+1} \circ \cdots \tau_{c_m,c_m+1}$ of σ into transpositions such that $3 < c_k$ for all $k \in [n]$. Define $u = X_{[c_1,c_1+1]} \cdot X_{[c_2,c_2+1]} \cdots X_{[c_m,c_m+1]}$. Then $[\![X_{[a,b]}]\!]_O^* = [\![\sigma]\!]_S = [\![u]\!]_O^*$. Then $X_{[a,b]} \sim_{\mathscr{R}_n^3} u$ by Lemma F.2. The proof follows by induction on m.

- **Base Case**. If $|u| = 0$, then $u \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot u$ by reflexivity.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|u| = k$, then $u \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot u$.
- **Inductive Step**. Assume that $m = k+1$. Define $v = X_{[c_1,c_1+1]} \cdot X_{[c_2,c_2+1]} \cdots X_{[c_k,c_k+1]}$. Then by definition $u = v \cdot X_{[c_m,c_m+1]}$. Since $c_m > 3$, then $u \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} v \cdot K_{[0,1,2,3]} \cdot X_{[c_m,c_m+1]}$ by the relation $X_{[c_m,c_m+1]} \cdot K_{[0,1,2,3]} \approx_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[c_m,c_m+1]}$. Since $|v| = k$, then by the inductive hypothesis $v \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot v$. Then $v \cdot K_{[0,1,2,3]} \cdot X_{[c_m,c_m+1]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot u$. Then by the transitivity of $(\sim_{\mathscr{R}_n^3})$, $u \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot u$ and the inductive step is established.

Then $u \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot u$ by the principle of induction. Since $X_{[a,b]} \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} u \cdot K_{[0,1,2,3]}$ and $K_{[0,1,2,3]} \cdot u \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[a,b]}$, then $X_{[a,b]} \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[a,b]}$ by the transitivity of $(\sim_{\mathscr{R}_n^3})$. □

Theorem F.15. *All instances of Relation (52) are derivable from $\mathscr{R}_n^3$.*

Proof. Let (c_0,c_1,c_2,c_3) an increasing sequence over $[n]$. Since (c_0,c_1,c_2,c_3) is increasing, then $k \leq c_k$ for $k \in [4]$. Then for each $k \in [4]$, define σ_k to be τ_{k,c_k} if $k \neq c_k$, or identity otherwise, and define $\sigma = \sigma_0 \circ \sigma_1 \circ \sigma_2 \circ \sigma_3$. Likewise, for each $k \in [4]$, define u_k to be $X_{[k,c_k]}$ if $k \neq c_k$, or ε otherwise, and let $u = u_0 \cdot u_1 \cdot u_2 \cdot u_3$. Clearly $\tau_{a,b} \circ \sigma = \sigma \circ \tau_{\sigma(a),\sigma(b)}$. Then,

$$[\![X_{[a,b]} \cdot u]\!]_O^* = [\![\tau_{a,b} \circ \sigma]\!]_S = [\![\sigma \circ \tau_{\sigma(a),\sigma(b)}]\!]_S = [\![u \cdot X_{[\sigma(a),\sigma(b)]}]\!]_O^*.$$

Then $X_{[a,b]} \cdot u \sim_{\mathscr{R}_n^3} u \cdot X_{[\sigma(a),\sigma(b)]}$ by Lemma F.2. Likewise, $X_{[\sigma(a),\sigma(b)]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} \overline{u} \cdot X_{[a,b]}$ by Lemma F.3. Since $a,b \notin \{c_0,c_1,c_2,c_3\}$, then $\sigma(a) > 3$ and $\sigma(b) > 3$. Then $X_{[\sigma a,\sigma b]} \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[\sigma a,\sigma b]}$ by Lemma F.14. Then the following derivation holds.

$$X_{[a,b]} \cdot u \cdot K_{[0,1,2,3]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot X_{[\sigma(a),\sigma(b)]} \cdot K_{[0,1,2,3]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot K_{[0,1,2,3]} \cdot X_{[\sigma(a),\sigma(b)]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} u \cdot K_{[0,1,2,3]} \cdot \overline{u} \cdot X_{[a,b]}$$

Since $\{a,b,c_0,c_1,c_2,c_3\}$ were arbitrary, then all instances of Relation (52) are derivable from $\mathscr{R}_n^3$. □

Lemma F.16. *Let (c_0,a,c_1,c_2,c_3) be an increasing sequence over $[n]$. For each $k \in [4]$, define σ_k to be τ_{k,c_k} if $k \neq c_k$, or identity otherwise. If $\sigma = \sigma_0 \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3$ and $\rho = \tau_{0,a} \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3$, then there exists an $\alpha \in \mathrm{S}(n)$ such that $\tau_{c_0,a} \circ \sigma = \rho \circ \alpha$ with α fixing $[4]$.*

Proof. By definition, $\rho(0) = a = \tau_{c_0,a}(\sigma(0))$, $\rho(1) = c_1 = \tau_{c_0,a}(\sigma(0))$, $\rho(1) = c_2 = \tau_{c_0,a}(\sigma(0))$, and $\rho(2) = c_2 = \tau_{c_0,a}(\sigma(0))$. Since $\mathrm{S}(n)$ is a group, then there exists an $\alpha \in \mathrm{S}(n)$ such that $\tau_{c_0,a} \circ \sigma = \rho \circ \alpha$. Assume that there exists a $k \in [4]$ such that $\alpha(k) \neq k$. Then $(\tau_{c_0,a} \circ \sigma)(\alpha(k)) \neq (\tau_{c_0,a} \circ \sigma)(k) = \rho(k)$. Then by contradiction, α fixes $[4]$. Then α decomposes into a sequence of transpositions over $[n] \setminus [4]$. □

Lemma F.17. *Let* $w = X_{[a_0,b_0]} \cdot X_{[a_1,b_1]} \cdots X_{[a_m,b_m]}$ *such that* $a_k > 3$ *and* $b_k > 3$ *for all* $k \in [m+1]$. *Then* $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$.

Proof. Let proof follows by induction on $|w|$.

- **Base Case**. If $|w| = 0$, then $w = \varepsilon$ and $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$ by the transitivity of $(\sim_{\mathscr{R}_n^3})$.
- **Inductive Hypothesis**. Assume that for some $k \in \mathbb{N}$, if $|w| = k$, then $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$
- **Inductive Step**. Assume that $|w| = k+1$. Define $v = X_{[a_0,b_0]} \cdot X_{[a_1,b_1]} \cdots X_{[a_k,b_k]}$. Then by definition $w = v \cdot X_{[a_m,b_m]}$. Since $a_m > 3$ and $b_m > 3$, then $X_{[a_m,b_m]} \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot X_{[a_m,b_m]}$ by Lemma F.14. As a result, $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} v \cdot K_{[0,1,2,3]} \cdot X_{[a_m,b_m]}$. Then by the inductive hypothesis, $v \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot v$. As a result, $v \cdot K_{[0,1,2,3]} \cdot X_{[a_m,b_m]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$. Then $w \cdot K_{[0,1,2,3]} \sim K_{[0,1,2,3]} \cdot w$ by the transitivity of $(\sim_{\mathscr{R}_n^3})$ and the inductive step is established.

Then by the principle of induction, $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$. □

Theorem F.18. *All instances of Relations (59), (60), (61) and (62) are derivable from* $\mathscr{R}_n^3$.

Proof. Let be (c_0, a, c_1, c_2, c_3) an increasing sequence. For each $k \in [4]$, define σ_k to be τ_{k,c_k} if $k \neq c_k$, or identity otherwise. Then define $\sigma = \sigma_0 \circ \sigma_1 \circ \sigma_2 \circ \sigma_3$ and $\rho = \tau_{0,a} \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3$. By Lemma F.16, there exists a sequence of transpositions $\alpha = \alpha_0 \circ \alpha_1 \circ \cdots \alpha_m$ such that $\tau_{c_0,a} \circ \sigma = \rho \circ \alpha$. Then define $w = [\![\alpha_0]\!]_S \cdot [\![\alpha_1]\!]_S \cdots [\![\alpha_m]\!]_S$. Next, for each $k \in [4]$, define u_k to be $X_{[k,c_k]}$ if $k \neq c_k$, or ε otherwise. Then define $u = u_0 \cdot u_1 \cdot u_2 \cdot u_3$, $v = X_{[c_0,a]} \cdot u_1 \cdot u_2 \cdot u_3$, and $w = [\![\alpha_0]\!]_S \cdot [\![\alpha_1]\!]_S \cdots [\![\alpha_m]\!]_S$. It follows that $[\![u]\!]_O^* = [\![\tau_{c_0,a} \circ \sigma]\!]_S = [\![\rho \circ \alpha]\!]_S = [\![v \cdot w]\!]$. Since u, v, and w are words over two-level operators of type X, then $X_{[c_0,a]} \cdot u \sim_{\mathscr{R}_n^3} v \cdot w$. Likewise, by Lemma F.3, $\overline{u} \cdot X_{[c_0,a]} \sim_{\mathscr{R}_n^3} \overline{w} \cdot \overline{v}$. Then $X_{[c_0,a]} \cdot u$ acts by conjugation on $K_{[0,1,2,3]}$ as follows.

$$X_{[c_0,a]} \cdot u \cdot K_{[0,1,2,3]} \cdot \overline{u} \cdot X_{[c_0,1]} \sim_{\mathscr{R}_n^3} v \cdot w \cdot K_{[0,1,2,3]} \overline{u} \cdot X_{[c_0,1]} \sim_{\mathscr{R}_n^3} v \cdot w \cdot K_{[0,1,2,3]} \cdot \overline{w} \cdot \overline{v}$$

Then by Lemma F.17, $w \cdot K_{[0,1,2,3]} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w$. Furthermore, $w \cdot \overline{w} \sim_{\mathscr{R}_n^3} \varepsilon$ Lemma F.3. Then w acts by conjugation on $K_{[0,1,2,3]}$ as follows.

$$w \cdot K_{[0,1,2,3]} \cdot \overline{w} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]} \cdot w \cdot \overline{w} \sim_{\mathscr{R}_n^3} K_{[0,1,2,3]}$$

Since $X_{[c_0,a]}{}^2 \sim_{\mathscr{R}_n^3} \varepsilon$ by Lemma F.2, then the following derivation also holds.

$$X_{[c_0,a]} \cdot u \cdot K_{[0,1,2,3]} \cdot \overline{u} \sim_{\mathscr{R}_n^3} X_{[c_0,a]} \cdot u \cdot K_{[0,1,2,3]} \cdot \overline{u} \cdot X_{[c_0,1]}{}^2 \sim_{\mathscr{R}_n^3} v \cdot w \cdot K_{[0,1,2,3]} \cdot \overline{w} \cdot \overline{v} \cdot X_{[c_0,1]} \sim_{\mathscr{R}_n^3} v \cdot K_{[0,1,2,3]} \cdot \overline{v} \cdot X_{[c_0,1]}$$

Since $\{a, b, c_0, c_1, c_2, c_3\}$ were all arbitrary, then all instances of Relation (59) holds. The cases of Relations (60), (61) and (62) follow symmetrically. □