Pauli and Clifford Groups

Ben OConnor

March 14, 2026

Contents

Introduction

These notes give an exposition of the group-theoretic structure underlying the Pauli and Clifford groups, with a focus on the ideas needed to understand the isomorphism between the projective Clifford group and the symplectic group over $\mathbb{F}_2$. The intended audience is researchers working in quantum circuit simulation who are mathematically literate but may not have a background in abstract algebra.

Sections 1–6 develop the necessary group theory from scratch: groups and subgroups, quotient groups, the isomorphism theorems, group presentations, abelian groups, and the commutator subgroup. These sections are terse by design and are not a substitute for a textbook treatment, but they establish the definitions and results that are applied in the later sections.

The remaining sections apply this machinery to the Pauli group $\mathcal{P}_n$, stabilizer states, and the Clifford group $\mathcal{C}_n$. The central result is that the projective Clifford group $\overline{\mathcal{C}}_n/\overline{\mathcal{P}}_n$ is isomorphic to the symplectic group $\mathrm{Sp}(2n, \mathbb{F}_2)$, and more broadly that the full projective Clifford group is a symplectic affine group. Along the way, we give an explicit inductive algorithm for decomposing Clifford gates into Hadamard, phase, and CNOT gates.

1 Group Basics

1.1 Core Definitions

Definition 1.1 (Group)**.** A *group* G is a set equipped with an associative binary operation $*\colon G \times G \to G$ such that

1. There exists an *identity* element $e \in G$:

$$e * g = g * e = g \text{ for all } g \in G$$

2. For each $g \in G$, g has an *inverse* g^{-1}:

$$g * g^{-1} = g^{-1} * g = e$$

In most cases, the group operator is only written explicitly when needed to clarify an ambiguity. For example, it is common to write gh instead of $g*h$ or $g \cdot h$.

Definition 1.2 (Subgroup)**.** A *subgroup* H of a group G is a subset $H \subset G$ that is a group under the binary operation on G restricted to H. It is common to write $H \leq G$ to specify that H is a subgroup of G.

Definition 1.3 (Subgroup Generated by A)**.** For any subset $A \subset G$, the *subgroup generated by* A is the smallest subgroup of G that contains A. Somewhat imprecisely, it is the group generated from A by (iteratively) taking products and inverses of all elements in A. More precisely, it is the intersection of all subgroups of G that contain A.

It is common to see a notation similar to $\langle g\colon \dots \rangle$ for the subgroup generated by the elements g that satisfy some condition. In some contexts it may be convenient to write $\langle A \rangle$ for the subgroup generated by A.

Definition 1.4 (Center)**.** The *center* $Z(G)$ of a group G is the collection of elements of G that commute with *all* elements of G:

$$Z(G) = \{g \in G\colon gh = hg \text{ for all } h \in G\}$$

It is simple to verify that $Z(G)$ is a subgroup of G. It is likewise simple to observe that the group operation for G is *commutative* if and only if $Z(G) = G$. Such a group is called a commutative group or an *abelian group* (named for Niels Henrik Abel).

Commutative groups are considerably simpler than non-commutative groups. A full classification of finite abelian groups is a standard topic of an undergraduate course in group theory. The full classification of the (non-commutative) finite *simple* groups was a monumental achievement of 20th century mathematics. The center of a group is one of many concepts that measures in some way the commutativity of a group. We will encounter others in what follows.

Definition 1.5 (Conjugation)**.** The *conjugate* g^h of g by h is $g^h = hgh^{-1}$.

It's possible to see the conjugate g^h defined as $h^{-1}gh$. Similar ambiguities arise for many definitions that involve a choice of sidedness. In most cases, the choice is largely inconsequential, as subsequent definitions often do not depend on this choice. (If they did, we might distinguish between a left- or right-sided conjugate.) For example, the definition of the centralizer below is identical for either choice of the definition of g^h.

If $g^h = g$, then we say that g is *centralized* by h. Notice that this is true if and only if $hg = gh$; that is, if g and h commute.

Definition 1.6 (Centralizer)**.** The *centralizer* $C_G(A)$ of A in G is the collection of elements of G that commute with all elements of A:

$$C_G(A) = \{g \in G \colon a^g = a \text{ for all } a \in A\}$$

It's simple to observe that $Z(G) = C_G(G)$, and $C_G(A)$ is always a subgroup of G. In the definition, it is not important whether or not A is subgroup or just any subset of G, but the centralizer remains the same if we instead consider the group *generated* by A. This is because any element that commutes with both a and b also commutes with their product (and their inverses).

Definition 1.7 (Commutator)**.** The (group) *commutator* $[g, h]$ of g and h is $[g, h] = ghg^{-1}h^{-1}$.

The commutator of two elements captures in some way the degree of non-commutativity between g and h. It is closely related to h^g, as $h^g h^{-1} = [g, h]$, but the commutator is more symmetric in its definition. We have that $[g, h] = e$ if and only if g and h commute, and $gh = [g, h]\, hg$. This second equation shows that the commutator of g and h can be thought of as a correction factor to the equation "$gh = hg$".

For later, we remark that $[g, h]^{-1} = [h, g]$, and $[g, h]^a = [g^a, h^a]$.

Definition 1.8 (Commutator Subgroup)**.** The *commutator subgroup* $G^{(1)}$ or $[G, G]$ (sometimes called the *derived subgroup*) of G is the group generated by all commutators of G:

$$[G, G] = \langle [g, h] : g, h \in G \rangle$$

The product of two commutators may fail to be a commutator, so it is necessary to take the group generated by all commutators. It is easy to see that G is commutative if and only if $[G, G] = \{e\}$. In this sense, the smaller the commutator subgroup is in G, the closer G is to being abelian.

Definition 1.9 (Group Homomorphism)**.** A *group homomorphism* $f : G \to H$ is a function f from a group G to a group H that respects the group operations on G and H. That is, if $*_G$ and $*_H$ are the operations on G and H, respectively, then for all $g_1, g_2 \in G$

$$f(g_1 *_G g_2) = f(g_1) *_H f(g_2)\,.$$

More succinctly, $f(g_1 g_2) = f(g_1)\, f(g_2)$.

If a homomorphism is also a bijection of sets, then it is called an *isomorphism*. If the homomorphism is a map from G to itself, then the terms homomorphism and isomorphism may be specified as endomorphism or automorphism, respectively. For later, we note that $\mathrm{Aut}(G)$, the set of automorphisms on G, is itself a group under function composition.

1.2 Exercises

Exercise 1.1. Show that $Z(G) \leq G$.

Exercise 1.2. Show that $C_G(A) \leq G$.

Exercise 1.3. Show that $Z(G) = C_G(G)$.

Exercise 1.4. Show that $g^h = g \iff gh = hg$.

Exercise 1.5. Show that $[g, h]^{-1} = [h, g]$.

Exercise 1.6. If $f \colon G \to H$ is a homomorphism, show that $f\big(g^{-1}\big) = f(g)^{-1}$ for all $g \in G$.

Exercise 1.7. If $f : G \to H$ is a homomorphism, show that the image of f is a subgroup of H; i.e., $\mathrm{Im}(f) = \{f(g) \mid g \in G\} \leq H$.

2 Quotient Groups

2.1 Equivalence Relations

Definition 2.1 (Equivalence Relation)**.** An equivalence relation on a set A is a relation $\sim$ that is

- Reflexive: $a \sim a$
- Symmetric: $a \sim b \implies b \sim a$
- Transitive: $a \sim b$ and $b \sim c \implies a \sim c$

Formally, a *relation* $\sim_R$ (not necessarily an equivalence relation) on a set A is any subset $R \subset A \times A$, and we define $a \sim_R b \iff (a, b) \in R$.

Definition 2.2 (Equivalence Class)**.** If A is a set with an equivalence relation $\sim$, then an *equivalence class* is a subset $C \subset A$ such that $a \sim b$ for all $a, b \in C$, and if $a \sim b$ for some $b \in C$, then $a \in C$.

It is straightforward to verify that if C_1 and C_2 are two equivalence classes for $(A, \sim)$, then either $C_1 = C_2$ or $C_1 \cap C_2 = \varnothing$. Moreover, for each $a \in A$, the set

$$C_a = \{x \in A \colon x \sim a\} \tag{2.1}$$

is an equivalence class with $a \in C_a$. We may call this the *class of* a and refer to any $a' \in C_a$ as a *representative of the class*.

Since each $a \in A$ is in exactly one equivalence class, we see that the set of equivalence classes forms a partition of A. That is, A is a disjoint union of its equivalence classes: $A = \cup_i C_i$ and $C_i \cap C_j \neq \varnothing$ if and only if $i = j$. We will use $\sqcup$ to denote disjoint union.

Conversely, any partition of a set $A = \sqcup_i C_i$ determines an equivalence relation by defining that $a \sim b$ if a and b are both in the same class C_i. The concepts of equivalence relations and partitions of a set are therefore identical.

Definition 2.3 (Quotients of Sets)**.** If A is a set with equivalence relation $\sim$, then the *quotient* of A by $\sim$ is the set of its equivalence classes:

$$A/\!\sim = \{C \colon C \text{ is an equivalence class for } \sim\}$$

The idea is that if $a \sim b$, then a and b are identified as the same element in the quotient $A/\!\sim$. In fact, there is always a natural map

$$q \colon A \to A/\!\sim \tag{2.2}$$

defined by $a \mapsto C_a$, and we have that $q(a) = q(b) \iff a \sim b$. We will often refer to this map as the quotient map, disambiguating which quotient map we are referring to as needed. We will typically refer to the image of a by the quotient map as $q(a)$, C_a, $[a]$, or $\bar{a}$, depending on the most convenient notation for the current context.

If $f \colon A \to B$ is any map, then the *pre-image* of $S \subset B$ is the set of elements of A that map to S:

$$f^{-1}(S) = \{a \in A \mid f(a) \in S\}$$

For $b \in B$, it is common to use the shorthand $f^{-1}(b) = f^{-1}(\{b\})$ for the pre-image of the element b. Considering the quotient map, for each $C \in A/\!\sim$ we have $q^{-1}(C) = C$. To interpret this clearly, we need to consider C as both an element of $A/\!\sim$ and a subset of A. We are saying that the pre-image of the element $C \in A/\!\sim$ is the set $C \subset A$.

In fact, this is really just another characterization of equivalence relations and partitions: they are equivalent to surjective maps. If $q \colon A \to B$ is *any* surjective map, then $A = \sqcup_{b \in B} q^{-1}(b)$. That is, A is partitioned by the pre-images of the elements of B, and this partition determines an equivalence relation on A. The conceptual difference is that we often start with a notion of equivalence, and this determines the (surjective) quotient map.

As a final point about quotients of sets, we consider maps out of quotients, $f \colon A/\!\sim \to B$. It can be inconvenient to work with the quotient object directly, so instead we often consider a map $F \colon A \to B$.

In fact, given f, the map F *always* exists by precomposing f with the quotient map:

$$F\colon A \to A/\!\sim \to B \tag{2.3}$$

That is, F factors as $F = f \circ q$.

Conversely, however, a map $F : A \to B$ doesn't always determine a map out of the quotient. F determines a map $f\colon A/\!\sim \to B$ if and only if $a_1 \sim a_2 \implies F(a_1) = F(a_2)$. In such a case, we might say that F descends to f, or that F *factors through* the quotient map:

$$\begin{array}{ccc} A & & \\ {\scriptstyle q}\downarrow & \searrow^{F} & \\ A/\!\sim & \xrightarrow[f]{} & B \end{array} \tag{2.4}$$

We record this as a theorem below.

Theorem 2.1 (Universal Property for Set Quotients)**.** *Let $F\colon A \to B$ be any map of sets, and suppose that A has an equivalence relation $\sim$. Then F factors through the quotient map (see eq.* (2.4) *above) if and only if $a_1 \sim a_2 \implies F(a_1) = F(a_2)$.*

The condition that $a_1 \sim a_2 \implies F(a_1) = F(a_2)$ means that the value of F does not depend on the choice of representative of the equivalence class. Therefore, F can be uniquely defined on the *class itself*, which is the content of the statement.

Everything about quotients of *sets* applies equally well to groups. But in order for quotient objects $G/\!\sim$ to be interesting as objects in the category of groups, we require that the quotient has a group operation that is compatible with the group operation on G. Specifically, we require that the quotient map $G \to G/\!\sim$ is a group homomorphism.

From here, it is a straightforward exercise in definitions to derive the precise properties needed of such an equivalence relation $\sim$ on a group. We skip the exercise and present its conclusions below.

2.2 Normal Subgroups

Definition 2.4 (Normalizer)**.** For a subgroup $H \leq G$, the *normalizer* $N_G(H)$ of H in G is the set of elements in G that fix H setwise:

$$N_G(H) = \{g \in G\colon h^g \in H \text{ for all } h \in H\}$$

If we define $H^g = gHg^{-1} = \left\{ghg^{-1}\colon h \in H\right\}$, then we can equivalently write

$$N_G(H) = \{g \in G\colon H^g = H\}$$

If $H^g = H$, we say that g *normalizes* H, or that H is normalized by g.

Unlike in the definition of the centralizer, it is important that H is a subgroup of G and not simply a subset. You *can* make the same definition for an arbitrary subset, but it's not generally useful or well-behaved, and in many cases will just be equal to the centralizer. Like the centralizer, however, any normalizer $N_G(H)$ is a subgroup of G, and $C_G(H) \leq N_G(H)$.

Definition 2.5 (Normal Subgroup)**.** A subgroup $N \leq G$ is a *normal subgroup* of G if $N_G(N) = G$; that is, if N is normalized by every element of G. We use $N \triangleleft G$ to denote that N is normal in G.

If $S \subset G$ is any subset, then we can define a relation on G by $g \sim_S h \iff g = sh$ for some $s \in S$. This relation is an equivalence relation precisely when the set S is a subgroup of G. We won't bother with all of the details, but basically there is a correspondence

- $e \in S \iff \sim_S$ is reflexive
- $\left(s \in S \implies s^{-1} \in S\right) \iff \sim_S$ is symmetric
- $(s_1, s_2 \in S \implies s_1 s_2 \in S) \iff \sim_S$ is transitive

As an example, suppose that $g \sim_S h$, and we want to show that $h \sim_S g$. Then we know that $g = sh$ for some $s \in S$, and we would like to show that $h = s'g$ for some $s' \in S$. Clearly, $h = s^{-1}g$, so we need S to have the property that $s \in S \implies s^{-1} \in S$.

For any subgroup $H \leq G$, we can thus define an equivalence relation $\sim_H$ and a quotient that we denote by G/H, the quotient of G by H, or G mod H.

This quotient still won't in general be a group in a way that is compatible with the quotient map $G \to G/H$. For that, we need H to be a *normal* subgroup of G.

Theorem 2.2 (Group Quotients are Quotients by Normal Subgroups)**.** *If $\sim$ is an equivalence relation on a group G such that $G \to G/\sim$ is a group homomorphism, then $\sim = \sim_N$ as defined above where N is a normal subgroup of G. Conversely, if $N \triangleleft G$, then $\sim_N$ is an equivalence relation on G, and G/N inherits a group structure from G such that the quotient map $G \to G/N$ is a group homomorphism.*

The group operation on the quotient $G/\sim$ is naturally determined by the group operation on G, since we are demanding that the quotient map be a group homomorphism:

$$\bar{g} * \bar{h} = \overline{g * h}$$

That is, the product of the class of g with the class of h is the class of the product gh. Implicitly, we have selected g and h to be representatives of their classes. If we were to choose different representatives, we would get the same result. This is a subtle but important point, and it is the reason why normality of N is required. The details are contained in the (omitted) proof of theorem 2.2.

We conclude with a basic but fundamentally important theorem about maps out of quotient groups.

Definition 2.6 (Kernel)**.** Let $f\colon G \to H$ be a group homomorphism. The *kernel* of f, $\ker(f)$, is the set of elements in G that map to the identity in H; i.e., the kernel of f is the the pre-image of e:

$$\ker(f) = \{g \in G \mid f(g) = e\} = f^{-1}(e)$$

Proposition 2.3 (Kernels are Normal Subgroups)**.** *Let $f : G \to H$ be a group homomorphism. Then* $\ker(f) \triangleleft G$.

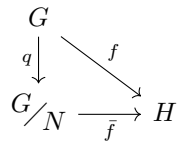
Proof Let $g \in G$ and $k \in \ker(f)$. Then

$$f\big(gkg^{-1}\big) = f(g)\, f(k)\, f\big(g^{-1}\big) = f(g)\, e f(g)^{-1} = e,$$

so $gkg^{-1} \in \ker(f)$, and $\ker(f) \triangleleft G$. □

Theorem 2.4 (Universal Property for Quotient Groups)**.** *Let $f : G \to H$ be a group homomorphism, and let $N \triangleleft G$. Then f factors through the quotient group $\bar{f} : G/N \to H$ if and only if $N \subset \ker(f)$. Moreover,* $\ker\big(\bar{f}\big) = \ker(f)/N$.

The theorem says we have the diagram below (in particular, the existence of the map $\bar{f}$) if and only if $N \leq \ker(f)$.



Proof Suppose that $f = \bar{f} \circ q$. If $k \in N = \ker(q)$, then

$$f(k) = \bar{f}(q(k)) = \bar{f}(e) = e,$$

so $k \in \ker(f)$ and $N \leq \ker(f)$

Now suppose that $N \subset \ker(f)$. For $g \in G$, we need show that $\bar{f}([g]) = \bar{f}(q(g)) = f(g)$ is a well-defined map. By theorem 2.1, $\bar{f}$ is well-defined if and only if $g_1 \sim_N g_2 \implies f(g_1) = f(g_2)$. So suppose that $g_1 \sim_N g_2$; i.e., $g_1 = ng_2$ for some $n \in N$. Then

$$f(g_2) = f(ng_1) = f(n)\, f(g_1) = e f(g_1) = f(g_1)\,.$$

It remains only to show that $\ker\big(\bar{f}\big) = \ker(f)/N$. Since $N \triangleleft G$, it immediately follows that $N \triangleleft \ker(f)$, and $\ker(f)/N = \{[k] \mid k \in \ker(f)\}$. Now we just observe that

$$[k] \in \ker\big(\bar{f}\big) \iff \bar{f}([k]) = f(k) = e \iff k \in \ker(f)\,,$$

so $\ker\big(\bar{f}\big) = \{[k] \mid k \in \ker(f)\} = \ker(f)/N$. □

2.3 Exercises

Exercise 2.1. Show that $N_G(A) \leq G$.

Example 2.2. The integers $\mathbb{Z}$ with its addition operation $+$ is an (abelian) group. (Considering the multiplication on $\mathbb{Z}$ as well makes the integers a *ring*, but for simplicity we can ignore this additional structure.) For any positive $n \in \mathbb{Z}$, the set $n\mathbb{Z} = \{n \cdot a \mid a \in \mathbb{Z}\}$ is a normal subgroup of Z, and the quotient group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ is the integers modulo n. For clarity, for $a \in \mathbb{Z}$ we will denote the class of a in $\mathbb{Z}_n$ by $[a]_n$.

Exercise 2.3. Show that the "map" from $\mathbb{Z}_3$ to $\mathbb{Z}_2$ "defined" by $[a]_3 \mapsto [a]_2$ is *not* well-defined.

Remark. This is an example of the fact that writing down a map out of a quotient object sometimes must be done with care. Somewhat subtly, a is being used as a representative of a class, but the map is only properly defined if the definition does not depend on which representative of the class is used. This is why theorems like theorem 2.4 help to correctly define maps out of quotients.

Exercise 2.4. Let A and B be sets with equivalence relations $\sim_A$ and $\sim_B$, respectively. Define an equivalence relation $\sim_{AB}$ on $A \times B$ by $(a_1, b_1) \sim_{AB} (a_2, b_2) \iff a_1 \sim_A a_2$ and $b_1 \sim_B b_2$. (You can check that this is in fact an equivalence relation.) Show that $[(a, b)] \mapsto ([a], [b])$ is a well-defined bijection from $A \times B/\sim_{AB}$ to $A/\sim_A \times B/\sim_B$.

Remark. The above exercise shows that the direct product of two quotients is a quotient of the direct product. This fact is used implicity in theorem 6.2.

Exercise 2.5. Let G_1 and G_2 be groups with $N_i \triangleleft G_i$. Show that $N_1 \times N_2 \triangleleft G_1 \times G_2$, and define an isomorphism between $G_1 \times G_2/N_1 \times N_2$ and $G_1/N_1 \times G_2/N_2$.

Exercise 2.6. For $N \leq G$, show that $\sim_N$ is an equivalence relation on G. (But don't forget that G/N only inherits a group structure from G when N is normal.)

Exercise 2.7. Let $\sim$ be an equivalence relation on G such that $q\colon G \to G/\sim$ is a homomorphism.

a) Show that the class of e is a normal subgroup of G.

b) Show that $\sim = \sim_N$ where N is the class of e.

Exercise 2.8. Let $m, n \in \mathbb{Z}$ be coprime positive integers. Show that the map $\mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$ defined by $[a]_{mn} \mapsto ([a]_m, [a]_n)$ is a well-defined isomorphism. You may assume the fact that if m and n are coprime, then if m and n both divide a, then mn divides a.

Exercise 2.9. Show that the map $\mathbb{Z}_{90} \to \mathbb{Z}_6 \times \mathbb{Z}_{15}$ defined by $[a]_{90} \mapsto ([a]_6, [a]_{15})$ is well-defined but not surjective.

Exercise 2.10. The image of a map $f\colon G \to H$ is $\mathrm{Im}(f) = \{f(g) \mid g \in G\}$. If f is a group homomorphism, show that $\mathrm{Im}(f) \leq G$.

3 Isomorphism Theorems

3.1 First Isomorphism Theorem

Groups can appear in many different circumstances, and given two groups it can be a challenge to understand the extent to which they share similar algebraic strcture. The strongest type of similarity between two structures would be an isomorphism, and identifying isomorphic structures is often a basic and essential part of any analysis of groups. By far the most foundational tool for carrying out any such analysis is the so-called First Isomorphism Theorem:

Theorem 3.1 (First Isomorphism Theorem)**.** *Let $f : G \to H$ be any group homomorphism. Then $G/\ker(f) \cong \mathrm{Im}(f)$.*

Proof By theorem 2.4, we can factor f as $f = \bar{f} \circ q$ where q is the natural quotient map $G \to G/\ker(f)$. Moreover, $\ker(\bar{f}) = \ker(f)/\ker(f) = \{e\}$, so $\bar{f}$ is injective. Trivially, f is surjective onto $\mathrm{Im}(f)$ and q is surjective onto $G/\ker(f)$. It follows that $\bar{f}$ is surjective onto $\mathrm{Im}(f)$, so $\bar{f}$ defines an isomorphism $G/\ker(f) \to \mathrm{Im}(f)$. □

As an immediate consequence, if $f\colon G \to H$ is a surjection, then $G/\ker(f) \cong H$.

3.2 Subgroup Lattice

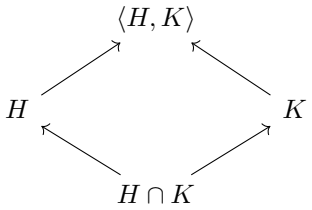
Since any $f : G \to H$ can be seen as an isomorphism $G/\ker(f) \cong H$, any homomorphism of groups requires that some quotient of the domain is isomorphic to some subgroup of the codomain. This basic observation suggests a strong connection between subgroups and homomorphisms, and as such an understanding of the subgroup structure of a group G is a powerful tool for any analysis of groups.

Remark. Any analysis of the subgroup structure of (finite) groups is woefully incomplete without a discussion of the Sylow theorems. But this section is already deviating from the main content, so we let this brief mention suffice as a "discussion".

Let $H, K \leq G$. It is not hard to see that $H \cap K \leq H, K$, and in fact $H \cap K$ is the largest subgroup contained in both H and K. Likewise, it is not hard to see that $H, K \leq \langle H, K\rangle = \langle g \mid g \in H \cup K\rangle$, and in fact $\langle H, K\rangle$ is the smallest sugroup containing both H and K.

Definition 3.1 (Subgroup Lattice)**.** Define $L(G) = \{H \mid H \leq G\}$ to be the set of subgroups of G. The operations $H \cap K$ and $\langle H, K\rangle$ give $L(G)$ the structure of a *lattice*. (In more general settings, these operations might be call the "meet" and "join", respectively.)

Our immediate focus will be on a local portion of the lattice:



Definition 3.2 (Partial Join)**.** For $H, K \leq G$, define $HK = \{hk \mid h \in H, k \in K\}$.

Always $H \cup K \subset HK \subset \langle H, K\rangle$, but it is not always the case that $HK = \langle H, K\rangle$, and in general HK is only a subset and not even a subgroup of G. The following proposition gives conditions under which $HK \leq G$.

Proposition 3.2 (Equality of Join and Partial Join)**.** *For $H, K \leq G$, the following are equivalent:*

1. $HK = KH$
2. $KH \subset HK$
3. $HK \leq G$
4. $HK = \langle H, K\rangle$

Proof Clearly, (1) implies (2).

Suppose that $KH \subset HK$. Let $hk, h'k' \in HK$. Since $kh' \in KH \subset HK$, we have $kh' = h''k'' \in HK$. So $hkh'k' = hh''k''k \in HK$. Since we also have $(hk)^{-1} = k^{-1}h^{-1} \in KH \subset HK$, this shows that $HK \leq G$.

Suppose that $HK \leq G$. We always have $H \cup K \subset HK \subset \langle H, K\rangle$. By definition of $\langle H, K\rangle$ as the group generated by $H \cup K$, it is the smallest subgroup that contains $H \cup K$. Since HK is a subgroup that contains $H \cup K$, we have $\langle H, K\rangle \leq HK$. So $HK = \langle H, K\rangle$

Suppose that $HK = \langle H, K\rangle$. Then $KH \subset \langle K, H\rangle = \langle H, K\rangle = HK$. Since HK is a subgroup, for any $hk \in HK$ we have $(hk)^{-1} \in HK$. Now $(hk)^{-1} = h_0k_0$, and

$$hk = (h_0k_0)^{-1} = k_0^{-1}h_0^{-1} \in KH.$$

So $HK \subset KH$ and $HK = KH$. □

Proposition 3.3. *If $H \leq N_G(K)$, then $HK \leq G$.*

Proof For any $hk \in HK$, $hk = \left(hkh^{-1}\right)h = k'h \in KH$. So $HK \subset KH$, which implies that $HK \leq G$. □

In general, the converse is not true. It may be the case that $HK \leq G$ with neither H normalizing K or K normalizing H.

Corollary 3.4. *Let $N \triangleleft G$. Then for any $H \leq G$, we have $HN = \langle H, N \rangle$.*

Proof $N \triangleleft G$ is equivalent to $N_G(N) = G$. So $H \leq G = N_G(N)$ implies that $HN \leq G$ and $HN = \langle H, N \rangle$. □

3.3 Diamond Isomorphism Theorem

Theorem 3.5 (Diamond Isomorphism Theorem)**.** *Let $H \leq N_G(K)$. Then $K \triangleleft HK$, $H \cap K \triangleleft H$ and*

$$HK/K \cong H/H \cap K.$$

Proof By assumption, H normalizes K, and a subgroup always normalizes itself so K normalizes K. It follows that the group generated by H and K, which is $\langle H, K \rangle$, normalizes K. So $K \triangleleft \langle H, K \rangle = HK$.

Let $g \in H \cap K$ and $h \in H$. Since H normalizes K and $g \in K$, we have $hgh^{-1} \in K$. Since $g \in H$, we also have $hgh^{-1} \in H$. So $hgh^{-1} \in H \cap K$, and $H \cap K \triangleleft H$.

To show the desired isomorphism, we make use of the first isomorphism theorem. Specifically, we will define a surjective map from H to HK/K with kernel $H \cap K$, and the result follows.

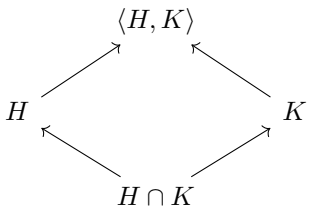
To begin, let $f : H \to HK/K$ be defined by the composition of the natural injection $\eta : H \to HK$ and the natural quotient $q : HK \to HK/K$. Since the kernel of q is K, $f(h) = q(\eta(h)) = q(h) = e$ if and only if $h \in K$. So the $\ker(f) = H \cap K$.

Now consider any $[hk] \in HK/K$. Since

$$[hk] = q(hk) = q(h)\, q(k) = q(h) = q(\eta(h)) = f(h)\,,$$

f is surjective. □

To summarize, we have been considering a portion of the lattice of the subgroups of G:



In general, $\langle H, K \rangle \neq HK$, and this is independent of whether or not $H \cap K \triangleleft H, K$. But $K \triangleleft \langle H, K \rangle \iff H \leq N_G(K)$, and either condition implies that $\langle H, K \rangle = HK$, $H \cap K \triangleleft H$, and $HK/K \cong H/H \cap K$.

Said differently, normality in a top edge of the diagram implies normality in the opposite (lower) edge and an isomorphism between the two quotients.

3.4 Lattice Isomorphism Theorem

Our final theorem concerns the full lattice of subgroups of G and how it relates to the lattice of subgroups of G/N for $N \triangleleft G$.

Theorem 3.6 (Lattice Isomorphism Theorem)**.** *Let $L(G \mid N) = \{H \leq G \mid N \leq H\}$. The map $H \mapsto H/N = \overline{H}$ gives a bijetion between $L(G \mid N)$ and $L(G/N)$. Moreover, if $q : G \to G/N$ is the quotient map, then $\overline{H} \mapsto q^{-1}(\overline{H})$ is inverse to $H \mapsto \overline{H}$, and the bijection respects normality and the lattice structure. Explicitly:*

1. *$H \triangleleft G \iff \overline{H} \triangleleft \overline{G}$ and $G/H \cong \overline{G}/\overline{H}$.*
2. *$H \leq K \iff \overline{H} \leq \overline{K}$*
3. *$H \cap K = \overline{H} \cap \overline{K}$*
4. *$\langle H, K \rangle = \langle \overline{H}, \overline{K} \rangle$*

3.5 Exercises

Exercise 3.1. Prove that $n\mathbb{Z} \subset m\mathbb{Z}$ if and only if m divides n.

Exercise 3.2. If m divides n, prove that $m\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_m/\mathbb{Z}_n \cong \mathbb{Z}_{\frac{n}{m}}$.

Exercise 3.3. Prove theorem 3.6. (Most parts of this result depend on little more than working through definitions. The isomorphism $G/H \cong \overline{G}/\overline{H}$ is perhaps the most interesting part of the proof, and even this is a standard application of the first isomorphism theorem.)

4 Group Presentations

4.1 Free Groups

Definition 4.1 (Free Group)**.** If S is any set, then the *free group* $F(G)$ on S is characterized by the following properties:

1. $S \subset F(S)$ with a natural inclusion of sets $\eta \colon S \to F(S)$
2. If H is any group and $f \colon S \to H$ is any map of sets, then there is a unique group homomorphism $F(f) \colon F(S) \to H$ such that $f = F(f) \circ \eta$

The second condition can be described by a commutative diagram:

$$\begin{array}{ccc} S & & \\ {\scriptstyle\eta}\downarrow & \searrow^{f} & \\ F(S) & \xrightarrow[F(f)]{} & H \end{array}$$

Note, however, that only the bottom map $F(f)$ is a group homomorphism. Both η and f are maps of sets, as S only has the structure of a set. The intuition is that S is the (set) data that fully determines the group $F(S)$, and it does so freely in the sense that it imposes no additional equalities on the group operation beyond those that are strictly required to be a group.

We won't go into the details of an explicit construction of $F(S)$, but essentially it is the set of all strings in S with concatentation as the group operation. The slight complexity is that we must add inverse elements and - in order to make it a group - we must be allowed to remove identity elements and adjacent inverse elements. Some explicit construction like this is necessary to prove the *existence* of the free group on S, but we emphasize that the properties given in definition 4.1 generally suffice to actually working with free groups.

4.2 Presentations

Let G be any group. We can always find some generators S of G. (The entire set $S = G$ generates G, for example, but typically we want to find a small generating set.) Now the map $f \colon S \to G$ then gives us a homomorphism $F(f) \colon F(S) \to G$, and since S generates G this map is necessarily surjective. By the first isomorphism theorem, $\overline{F(f)} \colon F(S) \,/\, \ker(F(f)) \to G$ is an isomorphism.

Definition 4.2 (Presentation of a Group)**.** Let G be a group. A *presentation* of G is a generating set $S \subset G$ and a set of relations $R \subset F(S)$ such that $N(R) = \ker(F(f))$ where $f \colon S \to G$ is the natural inclusion and $N(R)$ is the smallest normal subgroup of $F(S)$ containing R.

With these conditions,

$$\overline{F(f)} \colon {}^{F(S)}\!/_{N(R)} \to G$$

is an isomorphism.

Group presentations are notated in various ways, but the essential information is always the set of generators S and relations R. To simplify notation for this discussion, we will let $F(S \,|\, R)$ denote the presentation with generators S and relations R.

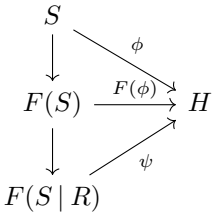
Given an element $w = s_1 \dots s_n \in F(S)$, a basic question is to ask whether or not $w = e$ in $F(S \mid R)$. (Since $u = v \iff uv^{-1} = e$, this is equivalent to asking whether or not two strings in $F(S)$ are equal in $F(S \mid R)$.) This is known as the word problem for the group G with presentation $F(S \mid R)$. For many common groups, algorithms exist to decide the word problem, but famously there exist groups with undecidable word problems, and the class of groups with decidable word problem is not uniformally decidable.

When an efficient algorithm for the word problem exists, it is often the case that any word $w \in F(S)$ can be reduced (using the relations in R) to some standard form $\sigma(w)$. The group operation on G can then be carried out by concatenating $\sigma(w_1)\,\sigma(w_2) = w'$ and reducing this back into standard form $\sigma(w')$. We will see some explicit examples of this with the Pauli group, which has a simple presentation and standard form.

We conclude with a sort of extension to part two of definition 4.1. This property of free groups ensured the existence of group homomorpisms $F(S) \to H$ for any set map $S \to H$. The following proposition considers what additional properties are needed of the set map $S \to H$ in order to generate a homomorphism $F(S|R) \to H$.

Proposition 4.1. *Let $F(S \mid R)$ be a presentation, and let H be any group. A map $\phi\colon S \to H$ induces a homomorpism $\psi\colon F(S \mid R) \to H$ if and only if $R \subset \ker(F(\phi))$.*

Proof This is an immediate consequence of theorem 2.4. We are considering a diagram of the form



Such a map ψ exists if and only if $N(R) \leq \ker(F(\phi))$, and since R normally generates $N(R)$, this is if and only if $R \subset \ker(F(\phi))$. □

As an immediate consequence, if G is a group with presentation $F(S \mid R)$, then we can specify homomorphisms out of G as any set of map of generators $S \to H$ such that each relation r maps to the identity in H.

5 Abelian Groups

5.1 Abelian Groups are Z-modules

Without going into too much detail, we record some basic properties of abelian groups. We will bold group elements $\mathbf{g}$ in this section to distinguish them from the integers $n \in \mathbb{Z}$ that will act on them. This notation is non-standard in this context and serves a primarily pedagogical purpose.

First, we note that the group operation in abelian groups is often written as $+$, and we typically write $\mathbf{e} = \mathbf{0}$ as a way to emphasize that the operation is commutative. With this notation, $\mathbf{g}^2$ becomes $\mathbf{g} + \mathbf{g}$, and we can write this as $2\mathbf{g}$. More generally, we can write $n\mathbf{g}$ for any $n \in \mathbb{Z}$. This action of $\mathbb{Z}$ on an abelian group satisfies several basic properties:

- $1\mathbf{g} = \mathbf{g}$
- $(n + m)\,\mathbf{g} = n\mathbf{g} + m\mathbf{g}$
- $(nm)\mathbf{g} = n(m\mathbf{g})$
- $n(\mathbf{g} + \mathbf{h}) = n\mathbf{g} + n\mathbf{h}$

Abelian groups therefore look a like a vector space over the integers. But the integers are not a field (multiplication is not invertible within $\mathbb{Z}$), so we cannot call it a vector space. Instead, such objects are known as $\mathbb{Z}$-modules.

Theorem 5.1. *Abelian groups are naturally $\mathbb{Z}$-modules.*

Now suppose that G is an abelian group and that for some $n \in \mathbb{N}$, we had that $n\mathbf{g} = \mathbf{0}$ for all $\mathbf{g} \in G$. Then we would have $(n+k)\mathbf{g} = n\mathbf{g} + k\mathbf{g} = 0 + k\mathbf{g} = k\mathbf{g}$. More generally, the value of $m\mathbf{g}$ only depends on the value of $m \pmod n$, and we can consider the action of $\mathbb{Z}$ on G as an action of $\mathbb{Z}_n$ on G. We can say in this case that G is a $\mathbb{Z}_n$-module.

When n is a prime (that we'll now call p), then the ring of intergers modulo p is actually a *field*, which we denote by $\mathbb{F}_p$. We therefore have the following theorem.

Theorem 5.2. *If $p \in \mathbb{Z}$ is a prime and G is an abelian group such that $p\mathbf{g} = 0$ for all $g \in G$, then G is an $\mathbb{F}_p$-module; that is, G is vector space over the field $\mathbb{F}_p$.*

5.2 Exercises

Exercise 5.1. For $a, n \in \mathbb{Z}$, if $\gcd(a,n) = d > 1$, show that there does not exist $k \in \mathbb{Z}$ such that $ak \equiv 1 \bmod n$.

Exercise 5.2. For $a, n \in \mathbb{Z}$ with $\gcd(a,n) = 1$, show that there exists $k \in \mathbb{Z}$ with $ak \equiv 1 \bmod n$. You may use the fact that there exists $k, p \in \mathbb{Z}$ such that $ak + pn = \gcd(a,n)$.

Exercise 5.3. If $p \in \mathbb{Z}$ is prime, show that for any non-zero $a \in \mathbb{Z}_p$, there is some $k \in \mathbb{Z}_p$ such that $ak = 1 \in \mathbb{Z}_p$.

6 Commutator Subgroup

6.1 Abelian Quotients

We now begin to connect some of the ideas from the previous sections as we move towards applications to the Pauli group. We begin with an important characterization of the commutator subgroup $[G,G]$.

Theorem 6.1 (Condition for Abelian Quotients)**.** *Let $N \triangleleft G$. Then G/N is abelian if and only if $[G,G] \leq N$.*

Proof Let $q\colon G \to G/N$ be the quotient map. Suppose that G/N is abelian. Then for each $g, h \in G$,

$$q([g,h]) = q\big(ghg^{-1}h^{-1}\big) = q(g)\,q(h)\,q(g)^{-1}\,q(h)^{-1} = [q(g)\,,q(h)] = e.$$

Thus, each commutator is an element of N. These elements generate $[G,G]$, so $[G,G] \leq N$.

Conversely, suppose that $[G,G] \leq N$. Then for each $\bar{g}, \bar{h} \in G/N$,

$$\big[\bar{g},\bar{h}\big] = \bar{g}\bar{h}\bar{g}^{-1}\bar{h}^{-1} = \overline{ghg^{-1}h^{-1}} = \overline{[g,h]} = e$$

where the last equality follows from the assumption that $[g,h] \in N$. Since all commutators on G/N equal e, this shows that G/N is abelian. □

6.2 Derived Commutator

We now prove a theorem that will be particular to our study of the Pauli and Clifford groups.

Theorem 6.2 (Derived Commutator)**.** *Let Z be the center of G. Then the map $[\,,]\colon G \times G \to [G,G]$ defined by $(g,h) \mapsto [g,h]$ factors through a map $[\langle\,,\rangle]\colon G/Z \times G/Z \to [G,G]$.*

$$\begin{array}{ccc} G \times G & & \\ \downarrow & \searrow^{[\,,]} & \\ G/Z \times G/Z & \xrightarrow[[\langle\,,\rangle]]{} & [G,G] \end{array}$$

Moreover, if $[G,G] \leq Z$, then $[\langle\,,\rangle]$ is a homomorphism of abelian groups in each factor.

Proof We first remark that we always have $Z \triangleleft G$. Now if $\bar{g}, \bar{h} \in G/Z$, we need to show that the map $\left[\langle \bar{g}, \bar{h} \rangle\right] \mapsto [g, h]$ is well-defined. That is, we need to show that if $g' \sim g$ and $h' \sim h$, then $[g, h] = [g', h']$. (This uses theorem 2.1 and exercise 2.4.)

First we observe that if $z \in Z$, then $[zg, h] = zgh\,(zg)^{-1}\,h^{-1} = ghg^{-1}h^{-1} = [g, h]$, since z commutes with all elements of G. Likewise, $[g, zh] = [g, h]$. Now if $g' \sim g$ and $h' \sim h$, then $g' = z_g g$ and $h' = z_h h$ for some some $z_g, z_h \in Z$. We therefore have $[g', h'] = [z_g g, z_h h] = [g, z_h h] = [g, h]$, as we wished to show.

Now we suppose that $[G, G] \leq Z$. First, notice that both G/Z and $[G, G]$ are abelian. We claim that

$$\left[\langle \bar{g}_1 + \bar{g}_2, \bar{h} \rangle\right] = \left[\langle \bar{g}_1, \bar{h} \rangle\right] + \left[\langle \bar{g}_2, \bar{h} \rangle\right] \tag{6.1}$$

and likewise for the second factor. This can be shown by

$$\left[\langle \bar{g}_1 + \bar{g}_2, \bar{h} \rangle\right] = [g_1 g_2, h] = g_1\,[g_2, h]\,g_1^{-1}\,[g_1, h] = [g_1, h]\,[g_2, h] = \left[\langle \bar{g}_1, \bar{h} \rangle\right] + \left[\langle \bar{g}_2, \bar{h} \rangle\right]$$

where the second equality can be checked directly and the third follows from the assumption that $[G, G] \leq Z$. □

We will refer to $[\langle\,,\,\rangle]$ as the derived commutator on G/Z. This map is *not* equal to the standard group commutator that is also defined on G/Z. For example, G may be non-abelian while G/Z is abelian. In this case, the commutator on G/Z is trivial, while the commutator on G is not. It is better to think of the derived commutator on G/Z as being in some sense the same as the commutator on G.

7 Pauli Group

7.1 Pauli Group $\mathcal{P}_1$

Definition 7.1 (Pauli Group)**.** The Pauli group $\mathcal{P}_1$ can be defined by the presentation

$$\mathcal{P}_1 = \langle X, Z, i \,|\, X^2 = Z^2 = i^4 = [i, X] = [i, Z] = e, [X, Z] = i^2 \rangle.$$

For notational convenience, we often idenfify $e = 1$ and $i^2 = -1$.

In our former notation, $\mathcal{P}_1 = F(S \,|\, R)$ with

$$S = \{X, Z, i\}$$

and

$$R = \left\{X^2, Z^2, i^4, iXi^{-1}X^{-1}, iZi^{-1}Z^{-1}, XZX^{-1}Z^{-1}i^{-2}\right\}.$$

The Pauli group typically appears as an explicit unitary representation where:

$$i = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} \qquad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

In this form, the Pauli group is a subgroup of the 2×2 unitary matrices $U(2)$. The 2×2 matrices over $\mathbb{C}$ also form a ring (actually, a $\mathbb{C}$-algebra) $\mathrm{Mat}^2(\mathbb{C})$, and occasionally we will use this additional structure. Of course, the Pauli group does not form a ring itself (the sum of two Pauli matrices need not be a Pauli matrix), but as a subset of $\mathrm{Mat}^2(\mathbb{C})$, the Pauli group has some useful properties. In particular, the Pauli group contains an orthogonal basis for $\mathrm{Mat}^2(\mathbb{C})$ as a vector space over $\mathbb{C}$.

The following statements are all easily verified and the proofs omitted.

Proposition 7.1. *For any $g, h \in \mathcal{P}_1$, $[g, h] = \pm 1$. In particular, $[\mathcal{P}_1, \mathcal{P}_1] = \{\pm 1\}$, which we can identify with $\mathbb{F}_2$ as an additive group.*

Proposition 7.2. *The center of the Pauli group is $Z(\mathcal{P}_1) = \langle i \rangle = \{\pm 1, \pm i\}$, and $\mathcal{P}_1/Z(\mathcal{P}_1) \cong \mathbb{F}_2^2$ as an $\mathbb{F}_2$-vector space.*

Proposition 7.3. *For any $g \in \mathcal{P}_1$, $g^2 = \pm 1$. In particular, $g^2 \in Z(\mathcal{P}_1)$.*

Observe that $[\mathcal{P}_1, \mathcal{P}_1] \leq Z(\mathcal{P}_1)$. The conditions of theorem 6.2 therefore apply, and the derived commutator

$$[\langle\,,\,\rangle] \colon \mathcal{P}_1/Z(\mathcal{P}_1) \times \mathcal{P}_1/Z(\mathcal{P}_1) \to [\mathcal{P}_1, \mathcal{P}_1] \tag{7.1}$$

is a bilinear map of $\mathbb{F}_2$-vector spaces.

7.2 Pauli Group $\mathcal{P}_n$

The group $\mathcal{P}_n$ is, roughly, the product of n copies of $\mathcal{P}_1$ glued along their centers. More precisely, it is the central product of n copies of $\mathcal{P}_1$ along their centers:

$$\mathcal{P}_n = \mathcal{P}_1^n \big/ N$$

where

$$N = \{(z_1, \ldots, z_n) \in Z(\mathcal{P}_1)^n \mid z_1 \cdots z_n = e\}.$$

The Pauli group $\mathcal{P}_n$ can also be defined by the presentation

$$\mathcal{P}_n = \left\langle s_1, \ldots, s_n, d_1, \ldots, d_n, i \mid s_j^2 = d_j^2 = i^4 = e, [s_j, d_j] = i^2, \text{ all other generators commute} \right\rangle.$$

We assume a basic familiarity with the Pauli group and its standard definition as a subgroup of $U(2^n)$, so we will not make heavy use of the somewhat abstract definitions above. For notational convenience we let $Z_n = Z(\mathcal{P}_n)$ and $\overline{\mathcal{P}}_n = \mathcal{P}_n / Z_n$.

A general element of $\mathcal{P}_n$ is

$$i^\alpha X_{(1)}^{\epsilon_1} Z_{(1)}^{\eta_1} \ldots X_{(n)}^{\epsilon_n} Z_{(n)}^{\eta_n},$$

and a general element of $\overline{\mathcal{P}}_n$ is

$$X_{(1)}^{\epsilon_1} Z_{(1)}^{\eta_1} \ldots X_{(n)}^{\epsilon_n} Z_{(n)}^{\eta_n}.$$

If we let $\hat{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ and $\hat{\eta} = (\eta_1, \ldots, \eta_n)$, then we can define

$$X^{\hat{\epsilon}} Z^{\hat{\eta}} = X_{(1)}^{\epsilon_1} \ldots X_{(n)}^{\epsilon_n} Z_{(1)}^{\eta_1} \ldots Z_{(n)}^{\eta_n} = X_{(1)}^{\epsilon_1} Z_{(1)}^{\eta_1} \ldots X_{(n)}^{\epsilon_n} Z_{(n)}^{\eta_n}.$$

With this notation, we can compute products in $\mathcal{P}_n$ as

$$i^{\alpha_1} X^{\hat{\epsilon}_1} Z^{\hat{\eta}_1} \cdot i^{\alpha_2} X^{\hat{\epsilon}_2} Z^{\hat{\eta}_2} = (-1)^{\hat{\eta}_1 \cdot \hat{\epsilon}_2} \cdot i^{\alpha_1 + \alpha_2} X^{\hat{\epsilon}_1 + \hat{\epsilon}_2} Z^{\hat{\eta}_1 + \hat{\eta}_2}$$

and commutators as

$$\left[i^{\alpha_1} X^{\hat{\epsilon}_1} Z^{\hat{\eta}_1}, i^{\alpha_2} X^{\hat{\epsilon}_2} Z^{\hat{\eta}_2} \right] = (-1)^{\hat{\eta}_1 \cdot \hat{\epsilon}_2 + \hat{\epsilon}_1 \cdot \hat{\eta}_2}.$$

With these eqaulties in mind, the following theorems are straightforward to check.

Theorem 7.4. *For any $g, h \in \mathcal{P}_n$, $[g, h] = \pm 1$. In particular, $[\mathcal{P}_n, \mathcal{P}_n] = \{\pm 1\}$, which we can identify with $\mathbb{F}_2$ as an additive group.*

Theorem 7.5. *The center of the Pauli group is $Z_n = \langle i \rangle = \{\pm 1, \pm i\}$, and $\overline{\mathcal{P}}_n \cong \overline{\mathcal{P}}_1^n \cong \mathbb{F}_2^{2n}$ as an $\mathbb{F}_2$-vector space. The isomorphism is given by the correspondence*

$$X^{\hat{\epsilon}} Z^{\hat{\eta}} \leftrightarrow (\hat{\epsilon}, \hat{\eta}).$$

Theorem 7.6. *For any $g \in \mathcal{P}_n$, $g^2 = \pm 1$. In particular, $g^2 \in Z_n$.*

As in the case $n = 1$, the conditions of Theorem 6.2 apply, and the map

$$[\langle\,,\,\rangle] \colon \overline{\mathcal{P}}_n \times \overline{\mathcal{P}}_n \to [\mathcal{P}_n, \mathcal{P}_n]$$

is an $\mathbb{F}_2$-bilinear map

$$[\langle\,,\,\rangle] \colon \mathbb{F}_2^{2n} \times \mathbb{F}_2^{2n} \to \mathbb{F}_2.$$

In fact, the following is true.

Theorem 7.7. *The map $[\langle\,,\rangle] \colon \overline{\mathcal{P}}_n \times \overline{\mathcal{P}}_n \to [\mathcal{P}_n, \mathcal{P}_n] \cong \mathbb{F}_2$ is a non-degenerate alternating form, so $\overline{\mathcal{P}}_n$ is a sympletic space over $\mathbb{F}_2$ of dimension $2n$.*

Proof This is nothing particular to the Pauli group. The form is alternating because $[g, g] = e$, and it is non-degenerate because $g \notin Z(G)$ implies the existence of $h \in G$ such that $[g, h] \notin e$. □

Symplectic spaces are well studied and much is known about their structure. I will write about them in another presentation, but for now we will make note of a few facts and definitions. A bilinear form $f : V \times V \to k$ is alternating if $f(v, v) = 0$ for all $v \in V$. It is non-degenerate if for each $v \in V$, there exists some $w \in V$ such that $f(v, w) \neq 0$.

If $v, w \in V$ and $f(v, w) \neq 0$, then $\{v, w\}$ is said to span a *hyperbolic subspace* of V. If H is a hyperbolic subspace of V, then we can always find an orthogonal complement to H; that is, $V = H \perp V'$ for some $V' \subset V$. Furthermore, we can always find an orthogonal decomposition of $V = H_1 \perp \cdots \perp H_n$. If $H'_1 \perp \cdots \perp H'_n$ is another such decomposition, then there is a form-preserving linear map $A : V \to V$ such that A maps each H_i isomorphically to each H'_i. We will show that the Clifford group is essentially all of the form-preserving maps, and this will be our primary characterization of the Clifford group.

7.3 Exercises

Exercise 7.1. Draw out the complete subgroup lattice for $\mathcal{P}_1$ and indicate which subgroups are normal.

Exercise 7.2. Show that the presentation given for $\mathcal{P}_1$ is isomorphic to the standard description as a unitary subgroup. You may assume that the matrices i, X, and Z generate this subgroup of order 16, as can be easily checked through direct computation.

Exercise 7.3.

1. Find a hyperbolic subspace of $\overline{\mathcal{P}}_3$ that contains $X^{(1,0,1)}Z^{(0,1,1)}$.
2. Find an orthogonal decomposition of $\overline{\mathcal{P}}_3$ that contains this hyperbolic subspace you found above.
3. Can you find every orthogonal decomposition of $\overline{\mathcal{P}}_3$ that extends $X^{(1,0,1)}Z^{(0,1,1)}$? How many are there?

8 Stabilizer States

8.1 Stabilizer-Destabilizer Pairs

Definition 8.1 (Stabilizer Set)**.** A subset $S = \{s_1, \ldots, s_n\} \subset \mathcal{P}_n$ is a *stabilizer set* if

- $s_i^2 = 1$ for each i
- $[s_i, s_j] = 1$ for each i, j
- $\{s_1, \ldots, s_n\}$ are distinct elements of $\overline{\mathcal{P}}_n$.

Definition 8.2 (Destabilizer Set)**.** Let $S = \{s_1, \ldots, s_n\} \subset \mathcal{P}_n$ be a stabilizer set. A subset $D = \{d_1, \ldots, d_n\} \subset \mathcal{P}_n$ is a *destabilizer set* for S if

- D is stabilizer set
- $[s_i, d_i] = -1$ for each i

The sets $S_Z = \left\{Z_{(1)}, \ldots, Z_{(n)}\right\}$ and $D_X = \left\{X_{(1)}, \ldots, X_{(n)}\right\}$ are an example of a stabilizer-destabilizer pair.

Theorem 8.1 (Existence of Destabilizer Sets)**.** *If $S \subset \mathcal{P}_n$ is a stabilizer set, then there exists a destabilizer set D for S. Moreover the pairs $\{s_i, d_i\}$ span a hyperbolic plane in $H_i \subset \overline{\mathcal{P}}_n$, and*

$$\overline{\mathcal{P}}_n = H_1 \perp \cdots \perp H_n.$$

Proof This is, essentially, a standard result about symplectic spaces. □

Corollary 8.2. *If (S, D) is a stabilizer-destabilizer pair, then $\{s_1, \ldots, s_n, d_1, \ldots, d_n\}$ is an $\mathbb{F}_2$-basis for $\overline{\mathcal{P}}_n$. Therefore, any element of $\mathcal{P}_n$ can be uniquely expressed as*

$$i^\alpha d_1^{\epsilon_1} s_1^{\eta_1} \cdots d_n^{\epsilon_n} s_n^{\eta_n}$$

for $\alpha \in \{0, 1, 2, 3\}$ and $\hat{\epsilon}, \hat{\eta} \in \{0, 1\}^n$. We denote such a representation as

$$i^\alpha D^{\hat{\epsilon}} S^{\hat{\eta}} = d_1^{\epsilon_1} \cdots d_n^{\epsilon_n} s_1^{\eta_1} \cdots s_n^{\eta_n} = d_1^{\epsilon_1} s_1^{\eta_1} \cdots d_n^{\epsilon_n} s_n^{\eta_n}$$

Let $s \in \mathcal{P}_n$ be an element of order 2. As an element of $U(2^n)$, s is a linear transformation of $\mathbb{C}^{2^n}$ with eigenvalues $\{\pm 1\}$. If we let $E_\alpha(s)$ be the eigenspace for s for eigenvalue α, then $\mathbb{C}^{2^n}$ decomposes as an orthogonal sum $\mathbb{C}^{2^n} = E_1(s) \perp E_{-1}(s)$. Moreover, we have the following important result.

Proposition 8.3. *Let $s \in \mathcal{P}_n$ be an element of order 2. If $d \in \mathcal{P}_n$ commutes with s, then d maps each $E_\alpha(S)$ isomorphically to itself. If d anti-commutes with s, then d maps $E_\alpha(S)$ isomorphically to $E_{-\alpha}(S)$.*

Proof Suppose that $d \in \mathcal{P}_n$ anti-commutes with s. Then for $v \in E_\alpha(s)$

$$sd \cdot v = -ds \cdot v = -\alpha d \cdot v.$$

So $d \cdot v \in E_{-\alpha}(s)$. Similarly, when d commutes with s we have $sd \cdot v = \alpha d \cdot v$. □

For notational consistency, it will be helpful to define $E^1(s) = E_{-1}(s)$ and $E^0(s) = E_1(s)$.

Definition 8.3. Let $S = \{s_1, \ldots, s_n\}$ be a stablizer set. Let $\hat{\epsilon} \in \{0,1\}^n$. Define the stabilizer eigenspace for S and $\hat{\epsilon}$ to be

$$E^{\hat{\epsilon}}(S) = \left\{ v \in \mathbb{C}^{2^n} \mid v \in E^{\epsilon_i}(s_i) \text{ for all } i \right\} = \bigcap_i E^{\epsilon_i}(s_i)\,.$$

For $\hat{0} = \{0, \ldots, 0\}$, any non-zero $v \in E^{\hat{0}}(S)$ is called a *stabilizer state* for S, and $sv = v$ for each $s \in S$.

Theorem 8.4. *Let (S, D) be a stabilizer-destabilizer pair. For all $\hat{\epsilon} \in \{0,1\}^n$, $\dim\big(E^{\hat{\epsilon}}(S)\big) = 1$. Moreover, if v_S is a stabilizer state for S, then for each $\hat{\epsilon}$*

$$D^{\hat{\epsilon}} \cdot v_S \in E^{\hat{\epsilon}}(S)\,.$$

Proof The second part of the statement is a consequence of proposition 8.3. The first part follows immediately from this. □

Corollary 8.5. *If (S, D) are a stabilizer-destabilizer pair and v_S is a stabilizer state, then*

$$\mathcal{B}(S, D) = \left\{ D^{\hat{\epsilon}} \cdot v_S \mid \hat{\epsilon} \in \{0,1\}^n \right\}$$

is an orthogonal basis of $\mathbb{C}^{2^n}$.

Theorem 8.6. *Let (S, D) be a stabilizer-destabilizer pair. Then*

$$\mathcal{B}_{\mathrm{Mat}}(S, D) = \left\{ D^{\hat{\epsilon}} S^{\hat{\eta}} \mid \hat{\epsilon}, \hat{\eta} \in \{0,1\}^n \right\}$$

is an orthogonal basis of $\mathrm{Mat}^{2^n}(\mathbb{C})$ with respect to the Hilbert-Schmidt inner product $\langle A, B \rangle = \mathrm{tr}(A^* B)$.

Proof The statement is easy to verify for $n = 1$. The general case follows from the equality $\mathrm{tr}(A_1 \otimes A_2) = \mathrm{tr}(A_1)\,\mathrm{tr}(A_2)$ and the fact that the trace is invariant under conjugation. □

We summarize the results as follows.

Theorem 8.7. *Let (S, D) be a stabilizer-destabilizer pair with v_S a stabilizer state for S. Then*

- *Any element of $\mathcal{P}_n$ can be uniquely expressed as $i^\alpha d_1^{\epsilon_1} s_1^{\eta_1} \cdots d_n^{\epsilon_n} s_n^{\eta_n}$ for $\alpha \in \{0,1,2,3\}$ and $\hat{\epsilon}, \hat{\eta} \in \{0,1\}^n$.*
- *(S, D) is an $\mathbb{F}_2$-basis for $\overline{\mathcal{P}}_n$, and the pairs $\{s_i, d_i\}$ determine a hyperbolic decomposition with respect to the symplectic form on $\overline{\mathcal{P}}_n$.*
- *$\mathcal{B}(S, D) = \left\{ D^{\hat{\epsilon}} \cdot v_S \mid \hat{\epsilon} \in \{0,1\}^n \right\}$ is an orthogonal basis for $\mathbb{C}^{2^n}$.*
- *$\mathcal{B}_{\mathrm{Mat}}(S, D) = \left\{ D^{\hat{\epsilon}} S^{\hat{\eta}} \mid \hat{\epsilon}, \hat{\eta} \in \{0,1\}^n \right\}$ is an orthogonal basis for $\mathrm{Mat}^{2^n}(\mathbb{C})$.*

8.2 Exercises

Exercise 8.1. Use the presentation of $\mathcal{P}_n$ to show that for any stabilizer-destabilizer pairs (S, D), (S', D'), there exists an automorphism $\phi\colon \mathcal{P}_n \to \mathcal{P}_n$ with $\phi(s_k) = s'_k$ and $\phi(d_k) = d'_k$ for all k.

9 Clifford Group

It is common to see the following definition of the Clifford group.

Definition 9.1 (Clifford Group)**.** The *Clifford group* $\mathcal{C}_n$ is the normalizer of $\mathcal{P}_n$ in $U(2^n)$. That is,

$$\mathcal{C}_n = N_{U(2^n)}(\mathcal{P}_n)\,.$$

This isn't what is actually meant, since often what is meant is a finite group, and $e^{i\theta}I$ is a central element of $U(2^n)$ so normalizes any subgroup. It's more accurate to say that $\mathcal{C}_n$ is generated by the non-diagonal elements of $U(2^n)$ that normalize $\mathcal{P}_n$. With this defintion, $\mathcal{C}_n$ contains 8 diagonal elements, the powers of $\zeta_8 = e^{i\pi/4}$.

Usually we want to quotient by the center of $\mathcal{C}_n$, so the distinction above is unimportant. What we care about is the action of $\mathcal{C}_n$ on $\mathcal{P}_n$ by conjugation, and central elements act trivially by conjugation. Our goal will be to make sense of the statement that the conjugation action of $\mathcal{C}_n$ on $\mathcal{P}_n$ determines an isomorphism

$$\overline{\mathcal{C}}_n \big/ \overline{\mathcal{P}}_n \cong \mathrm{Sp}(2n, \mathbb{F}_2). \tag{9.1}$$

Definition 9.2 (Automorphism Group)**.** For any group G, its *automorphism group* $\mathrm{Aut}(G)$ is the set of automorphisms

$$\{f : G \to G \mid f \text{ is an isomorphism}\} \tag{9.2}$$

with group operation given by function composition.

Definition 9.3 (Conjugation Map)**.** If $N \triangleleft G$, then the conjugation map $c : G \to \mathrm{Aut}(N)$ is defined by

$$c(g)(n) = n^g.$$

That is, $c(g)$ is the automorphism of N given by conjugation by g.

Definition 9.4 (Inner Automorphism Group)**.** The group $\mathrm{Inn}(G)$ of *inner automorphisms* of a group G is the image of G in $\mathrm{Aut}(G)$ under the conjugation map. The kernel of this map is always the center of G, so there is an isomorphism

$$G \big/ Z(G) \cong \mathrm{Inn}(G).$$

It is always true that $\mathrm{Inn}(G) \triangleleft \mathrm{Aut}(G)$, and the group $\mathrm{Out}(G)$ of *outer automorphisms* of G is

$$\mathrm{Out}(G) = \mathrm{Aut}(G) \big/ \mathrm{Inn}(G).$$

Considering the Pauli group, there is special automorphism that swaps i and $-i$; essentially a complex conjugation of the group. This cannot be achieved by a conjugation action within $U(2^n)$, so we consider a restricted set of automorphisms that fix i

$$\mathrm{Aut}_0(\mathcal{P}_n) = \{f \in \mathrm{Aut}(\mathcal{P}_n) \mid f(i) = i\}\,.$$

Similarly, we'll define $\mathrm{Out}_0(\mathcal{P}_n) = \mathrm{Aut}_0(\mathcal{P}_n)/\mathrm{Inn}(\mathcal{P}_n)$.

Theorem 9.1. *Let $\phi \in \ker(c)$, where c is the conjugation map $c : \mathcal{C}_n \to \mathrm{Aut}_0(\mathcal{P}_n)$. Then ϕ is diagonal; i.e., $\phi \in Z(\mathcal{C}_n)$ and $\ker(c) = Z(\mathcal{C}_n)$.*

Proof For $\phi \in \ker(c)$, ϕ is a stabilizer of $\mathcal{P}_n$. Since $\mathcal{P}_n$ contains a basis of $\mathrm{Mat}^{2^n}(\mathbb{C})$, anything that stabilizes $\mathcal{P}_n$ commutes with any element of $\mathrm{Mat}^{2^n}(\mathbb{C})$. These are well known to be the diagonal matrices. □

Theorem 9.2. *Let $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$. Then $\phi = c(g)$ for some $g \in \mathcal{C}_n$. In other words, the conjugation map $c : \mathcal{C}_n \to \mathrm{Aut}_0(\mathcal{P}_n)$ is surjective.*

This theorem is the most computationally intensive, and we save the proof for later. We will give an explicit inductive algorithm for determining g using only the Hadamard, phase and CNOT gates, which will further prove that $\mathcal{C}_n$ is generated by $\mathcal{P}_n$ along with these three types of non-Pauli Clifford gates. We also note that we have proven $\mathrm{Aut}_0(\mathcal{P}_n) \cong \overline{\mathcal{C}}_n$.

Theorem 9.3. *Let $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$, and let (S, D) be a stabilizer-destabilizer pair. Suppose that $\phi(s_i) = (-1)^{\epsilon_i} s_i$ and $\phi(d_i) = (-1)^{\eta_i} d_i$ for all i. Then $\phi = c\big(D^{\hat{\epsilon}} S^{\hat{\eta}}\big)$. In particular, $\phi \in \mathrm{Inn}(\mathcal{P}_n)$.*

Proof Since $\{s_i, d_i\}$ are the only non-commuting elements, we see that

$$c\big(D^{\hat{\epsilon}} S^{\hat{\eta}}\big)(s_i) = D^{\hat{\epsilon}} S^{\hat{\eta}} s_i S^{\hat{\eta}} D^{\hat{\epsilon}} = d_i^{\epsilon_i} s_i d_i^{\epsilon_i} = (-1)^{\epsilon_i} s_i \tag{9.3}$$

and similarly for d_i. □

Definition 9.5 (Characteristic Subgroup)**.** A subgroup $H \leq G$ is called a *characteristic subgroup* of G if $\phi(H) = H$ for all $\phi \in \mathrm{Aut}(G)$.

The center and commutator subgroups of G are easily seen to be characteristic, essentially because $\phi([g, h]) = [\phi(g), \phi(h)]$ for any automorphism ϕ.

Theorem 9.4. *If H is characteristic is G, then there is a natural map $\mathrm{Aut}(G) \to \mathrm{Aut}(G/H)$ taking an automorphism $\phi \in \mathrm{Aut}(G)$ to $\bar{\phi}$ as in the diagram below.*

$$\begin{array}{ccc} G & \xrightarrow{\ \phi\ } & G \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q} \\ G/H & \xrightarrow[\ \bar{\phi}\ ]{} & G/H \end{array}$$

Proof The crux of the argument is the existence of $\bar{\phi}$, which exists precisely when H is in the kernel of $q \circ \phi$. We can compute this kernel

$$\ker(q \circ \phi) = \phi^{-1}\big(q^{-1}(e)\big) = \phi^{-1}(H) = H$$

where the last inequality is a consequence of the assumption that H is characteristic in G. □

Applying this to the Pauli group, we get a map $Q : \mathrm{Aut}_0(\mathcal{P}_n) \to \mathrm{Aut}\big(\overline{\mathcal{P}}_n\big)$. Since $\overline{\mathcal{P}}_n = \mathbb{F}_2^{2n}$, its automorphism group is $\mathrm{GL}(2n, \mathbb{F}_2)$, the group of invertible $2n \times 2n$ matrices with elements in $\mathbb{F}_2$. But we can say more about the image of Q.

Theorem 9.5. *Let $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$. Then $Q(\phi) \in \mathrm{Sp}(2n, \mathbb{F}_2)$.*

Proof First, we note that $\phi([g, h]) = [g, h]$ for all $g, h \in \mathcal{P}_n$. This follows from the fact that ϕ acts automorphically on the characteristic subgroup $[\mathcal{P}_n, \mathcal{P}_n] = \mathbb{F}_2$, and the only automorphism of $\mathbb{F}_2$ is the identity map. Now we can compute

$$\big[\big\langle Q(\phi)(\bar{g}), Q(\phi)(\bar{h})\big\rangle\big] = \Big[\Big\langle \overline{\phi(g)}, \overline{\phi(h)}\Big\rangle\Big] = [\phi(g), \phi(h)] = \phi([g, h]) = [g, h].$$

So $Q(\phi)$ preserves the symplectic form on $\mathcal{P}_n$; i.e. $Q(\phi) \in \mathrm{Sp}(2n, \mathbb{F}_2)$. □

Theorem 9.6. *Let $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$. Then $\phi \in \ker(Q)$ if and only if $\phi \in \mathrm{Inn}(\mathcal{P}_n)$.*

Proof If $\phi \in \mathrm{Inn}(\mathcal{P}_n)$, then ϕ is conjugation by some $g \in \mathcal{P}_n$. Conjugating by an element of $\mathcal{P}_n$ only changes the sign of elements, so the action is trivial on $\overline{\mathcal{P}}_n$ and $\phi \in \ker(Q)$.

If $\phi \in \ker(Q)$, then for all $h \in \mathcal{P}_n$, $\phi(g) = z_g g$ for some $z_g \in Z_n$. By Theorem 9.3, $\phi \in \mathrm{Inn}(\mathcal{P}_n)$. □

Putting this altogether, we have that $\mathrm{Im}(Q) \cong \mathrm{Aut}_0(\mathcal{P}_n)/\ker(Q) \cong \overline{\mathcal{C}}_n/\overline{\mathcal{P}}_n$, and this is a subgroup of $\mathrm{Sp}(2n, \mathbb{F}_2)$. So the only remaining thing to show is that Q surjects onto $\mathrm{Sp}(2n, \mathbb{F}_2)$.

Theorem 9.7. *For any $M \in \mathrm{Sp}(2n, \mathbb{F}_2)$, there exists some $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$ with $Q(\phi) = M$.*

Proof Consider any stabilizer-destabilizer pair (S, D) in $\mathcal{P}_n$ and its image $(\bar{S}, \bar{D})$ in $\overline{\mathcal{P}}_n$. Then $(\bar{S}, \bar{D})$ forms a symplectic basis for $\overline{\mathcal{P}}_n$. Since $M \in \mathrm{Sp}(2n, \mathbb{F}_2)$, the image of $(\bar{S}, \bar{D})$ in $\overline{\mathcal{P}}_n$ is also a symplectic basis $(\bar{S}', \bar{D}')$.

Lifting this back to $\mathcal{P}_n$, we get a new set (S', D') that is also a stabilizer-destabilizer pair (possibly after multiplying some elements by i to ensure that they have order 2). By exercise 8.1, there is an automorphsim ϕ that takes (S, D) to (S', D'). $Q(\phi)$ therefore maps $(\bar{S}, \bar{D})$ to $(\bar{S}', \bar{D}')$, and since $(\bar{S}, \bar{D})$ is a basis of $\overline{\mathcal{P}}_n$, we conclude that $Q(\phi) = M$. □

10 Change of Pauli Basis

In this section we aim to prove theorem 9.2. To do so, we will need only three types of elements from the Clifford group.

Definition 10.1. The following elements are the Hadamard, phase, and CNOT gates, respectively:

1. $H_j = \frac{1}{\sqrt{2}}(Z_j + X_j)$
2. $P_j = \sqrt{Z_j} = \alpha I + \bar{\alpha} Z_j$ where $\alpha = \frac{1+i}{2}$
3. $C_{jk} = \frac{1}{2}(I + Z_j + X_k - Z_j X_k)$

There is much more than can be said about these gates, but for our purposes all that we will need are the following conjugation identities.

Proposition 10.1. *On the Pauli generators $\{Z_1, \ldots, Z_n, X_1, \ldots, X_n\}$, we have the following identities:*

1. $X_j^{H_j} = Z_j$
2. $Z_j^{H_j} = X_j$
3. $X_j^{P_j} = iX_jZ_j$
4. $Z_k^{C_{jk}} = Z_jZ_k$
5. $X_j^{C_{jk}} = X_jX_k$

All other Pauli generators commute with H_j, P_j, and C_{jk}.

Our goal is to show that given any $\phi \in \mathrm{Aut}_0(\mathcal{P}_n)$, we can find some $g \in \mathcal{C}_n$ such that the automorphism ϕ is conjugation by g. For $\mathcal{P}_1$, this is an easy computation using only H and P. For the general case, the algorithm proceeds inductively: we find a Clifford element g that conjugates $\phi(Z_1)$ to Z_1 and $\phi(X_1)$ to X_1, then observe that the remaining generators are confined to the subgroup $\mathcal{P}_{n-1}$ generated by $\{X_2, \ldots, X_n, Z_2, \ldots, Z_n\}$.

The core of the algorithm is the reduction of $\phi(Z_1)$ and $\phi(X_1)$ to Z_1 and X_1. Working in the $\mathbb{F}_2$-vector representation $(\hat{\epsilon} \mid \hat{\eta})$ of $\overline{\mathcal{P}}_n$, conjugation by each gate type acts as follows:

- H_j: swaps $\epsilon_j \leftrightarrow \eta_j$
- P_j: $\eta_j \mapsto \eta_j + \epsilon_j$
- C_{jk}: $\epsilon_k \mapsto \epsilon_k + \epsilon_j$ and $\eta_j \mapsto \eta_j + \eta_k$

All other coordinates are unchanged. To simplify the discussion, we ignore factors of i; the procedure will at most produce incorrect signs, which can be corrected by conjugating with an element of $\mathcal{P}_n$.

Phase 1: $\phi(Z_1) = (\hat{\epsilon} \mid \hat{\eta}) \mapsto Z_1 = \big(\hat{0} \mid e_1\big)$**.** The sequence of reductions is:

$$\begin{aligned}(\hat{\epsilon} \mid \hat{\eta}) &\xrightarrow{\text{ensure some } \eta_j=1} (\hat{\epsilon} \mid \hat{\eta}') \text{ with } \eta'_j = 1 \\ &\xrightarrow{\text{move to position 1}} (\hat{\epsilon} \mid \hat{\eta}'') \text{ with } \eta''_1 = 1 \\ &\xrightarrow{(\epsilon_j,\eta_j)\to(0,*) \text{ for all } j} \big(\hat{0} \mid \hat{\eta}'''\big) \text{ with } \eta'''_1 = 1 \\ &\xrightarrow{\eta_j \to 0 \text{ for } j>1} \big(\hat{0} \mid e_1\big) = Z_1\end{aligned}$$

Phase 2: $\phi(X_1)^g = (\hat{\epsilon} \mid \hat{\eta}) \mapsto X_1 = \big(e_1 \mid \hat{0}\big)$**, preserving** Z_1**.** Here g is the Clifford element produced by Phase 1. Since $\phi(X_1)^g$ anti-commutes with $Z_1 = \phi(Z_1)^g$, necessarily $\epsilon_1 = 1$. The reductions are dual to Phase 1:

$$\begin{aligned}(\hat{\epsilon} \mid \hat{\eta}) \text{ with } \epsilon_1 = 1 &\xrightarrow{(\epsilon_j,\eta_j)\to(*,0) \text{ for all } j} \big(\hat{\epsilon}' \mid \hat{0}\big) \text{ with } \epsilon'_1 = 1 \\ &\xrightarrow{\epsilon_j \to 0 \text{ for } j>1} \big(e_1 \mid \hat{0}\big) = X_1\end{aligned}$$

Each operation in Phase 2 must preserve Z_1, which constrains the available gates; the details are given below.

Let $\phi(Z_1) = X_{(1)}^{\epsilon_1} Z_{(1)}^{\eta_1} \dots X_{(n)}^{\epsilon_n} Z_{(n)}^{\eta_n}$. Suppose first that all $\eta_i = 0$. Then some $\epsilon_j = 1$. Conjugation by H_j then sets $\eta_j = 1$.

So we may now assume that some $\eta_j = 1$. If $j \neq 1$, then conjugation by C_{1j} sets $\eta_1 = 1$.

Now consider each j with $\epsilon_j = 1$. If $\eta_j = 0$, then conjugation by H_j sets $\epsilon_j = 0$ and $\eta_j = 1$. If $\eta_j = 1$, then conjugation by P_j followed by H_j sets $\epsilon_j = 0$. In this way, we have all $\epsilon_j = 0$ and $\eta_1 = 1$.

Finally, for each $j > 1$ with $\eta_j = 1$, conjugation by C_{j1} sets $\eta_j = 0$. We have thus found a conjugation procedure taking $\phi(Z_1) \mapsto Z_1$.

After this conjugation, $\phi(X_1)$ will equal some $X_{(1)}^{\epsilon_1} Z_{(1)}^{\eta_1} \dots X_{(n)}^{\epsilon_n} Z_{(n)}^{\eta_n}$. Necessarily, $\epsilon_1 = 1$, since this element anti-commutes with Z_1. We can repeat a similar procedure as above, but with all of the following conjugations preserving Z_1.

Consider all $j > 1$ with $\eta_j = 1$. If $\epsilon_j = 0$, then conjugation by H_j sets $\eta_j = 0$ and $\epsilon_j = 1$. If $\epsilon_j = 1$, then conjugation by P_j sets $\eta_j = 0$. In this way, we have all $\eta_j = 0$ and $\epsilon_1 = 1$.

Finally, for each $j > 1$ with $\epsilon_j = 1$, conjugation by C_{1j} sets $\epsilon_j = 0$. Since each $\eta_j = 0$, these conjugations preserve Z_1 as required.

This procedure produces a conjugation mapping $\phi(Z_1) \mapsto Z_1$ and $\phi(X_1) \mapsto X_1$. It remains to show that this reduces the problem to $\mathcal{P}_{n-1}$. Since $Z_1 = \phi(Z_1)^g$ and $X_1 = \phi(X_1)^g$ commute with $\phi(Z_j)^g$ and $\phi(X_j)^g$ for $j \geq 2$, the images $\phi(Z_j)^g$ and $\phi(X_j)^g$ must lie in the subgroup generated by $\{X_2, \dots, X_n, Z_2, \dots, Z_n\}$, which is precisely the centralizer of $\{Z_1, X_1\}$ in $\mathcal{P}_n$. This subgroup is isomorphic to $\mathcal{P}_{n-1}$, and inductively we find $h \in \mathcal{C}_{n-1}$ such that $\left(\phi(Z_j)^g\right)^h = Z_j$ and $\left(\phi(X_j)^g\right)^h = X_j$ for $j \geq 2$. We then have $Z_j^{(hg)^{-1}} = \phi(Z_j)$ and $X_j^{(hg)^{-1}} = \phi(X_j)$ for all j, so $\phi = c\big((hg)^{-1}\big)$.

Remark. It is worth pointing out one fact, which in some sense is the essential feature of the Clifford group. In the above procedure, Z and X could have been exchanged for any stabilizer-destabilizer pair. Then H, P, and C could have been defined relative to this new generating set, and everything would proceed identically. The Pauli and Clifford groups are entirely symmetric around stabilizer-destabilizer pairs. Essentially anything that is true for the standard computational basis is equally true for a general stabilizer-destabilizer pair.