# INTRODUCTION TO MATHEMATICAL QUANTUM ERROR CORRECTION

DHEERAN E. WIGGINS

*Additional Facilitator: Micah E. Fogel*

Abstract. Quantum computation (QC) and information (QI) are relatively young fields, conceived less than half a century ago. In the subsequent decades, those fluent in computer science, mathematics, or physics have seen an increased demand for QC and QI talent across industry and academia. Still, quantum computation retains an inherent constraint–"quantum noise"–a susceptibility to errors caused by quantum mechanical phenomena. Thus, the study of Quantum Error Correction (QEC) was born, giving way to a series of mathematical, algorithmic, and physical techniques used to identify and correct any such errors. After taking care of some group theory and linear algebra preliminaries, we develop some major tools of QEC, such as the Knill-Laflamme subspace condition and the stabilizer formalism, via an abstract, mathematical perspective. We then discuss some contemporary research results, arriving at Operator Quantum Error Correction (OQEC), developed by Poulin et al. in 2006.

## Contents

1. Preliminaries 2
1.1. Setting and Structure 2
1.2. Naive Set Theory 3
1.3. From Sets to Groups: Adding an Operation 5
1.4. From Groups to Vector Spaces: Adding Complex Linearity 8
2. The Subspace Condition 14
2.1. Tensor Products and Quantum Channels 14
2.2. The Kraus Representation 19
2.3. Anticliques and the Knill-Laflamme Condition 21
3. The Stabilizer Formalism 22
3.1. An Elementary Code 22
3.2. Stabilizer Groups, Spaces, and Codes 23
4. Operator Quantum Error Correction 24
4.1. Noiseless Subsystems 25
4.2. Revisitng the Stabilizer Formalism 26
Acknowledgements 28
References 29

*Key words and phrases.* Hilbert Spaces, Quantum Channels, Knill-Laflamme, Stabilizer Formalism.

These lecture notes were compiled for the 2025 Intersession at the Illinois Mathematics and Science Academy (IMSA). Lectures ($\approx$ **20** hours) were presented to secondary students over five days. There are minimal prerequisites to read these notes, though a moderate comfortability with mathematical abstraction or quantum mechanics may complement them nicely.

## 1. Preliminaries

Before jumping into quantum error correction, I want to take some time to contextualize the subject of quantum information. I will also give a brief, proof-less review of sets, groups, vector spaces, and tensor products, so our tool belt is ready for the incoming mathematics. If you are not already familiar with the material, I encourage you to return frequently to this review while reading the later sections.

### 1.1. Setting and Structure.

Let us take our physical system to be a collection of qubits, their environment, and any forces acting between the qubits. Then, per quantum mechanics, we may encode (Fig. 1) the data of this system in a *Hilbert space* $\mathcal{H}$.

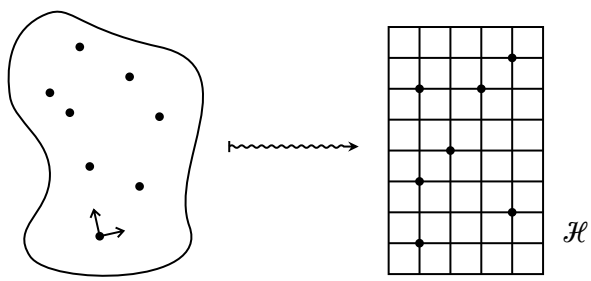


Figure 1. The modeling of a physical system as a Hilbert space $\mathcal{H}$.

Contemporary quantum mechanics follows four axioms describing the way to connect physical reality to our Hilbert space mathematics. We will state the axioms now, returning to some of them after our mathematical preliminaries.[1]

**Axiom 1.1** (State Space)**.** Any quantum system $Q$ is represented by a complex Hilbert space $\mathcal{H}^Q$, called the state space. States of the system are represented by unit-trace, positive semi-definite operators acting on $\mathcal{H}$, called density operators.

**Axiom 1.2** (Multiple System)**.** Any pair of join quantum systems $A, B$ can be represented by a tensor product Hilbert space

$$\mathcal{H}^{AB} := \mathcal{H}^A \otimes \mathcal{H}^B.$$

**Axiom 1.3** (System Evolution)**.** A quantum system $Q$ undergoing closed evolution is described by a unitary transformation on the state space $\mathcal{H}^Q$.

**Axiom 1.4** (Measurement)**.** Every measurement of a finite dimensional quantum system is described by a set of orthogonal projectors $\{P_i\}_{i=1}^r$ such that $\sum_{i=1}^r P_i = I^Q$. If $\rho$ is the state of $Q$ prior to measurment, then with probability $\mathbb{P}(i) = \mathrm{tr}(P_i\rho)$, the post-measurement state will be

$$\rho_i = \frac{P_i \rho P_i}{p(i)}.$$

[1]The first three axioms will be most important in our development of quantum channels and error correction. Measurement is a process which extracts classical information, like bits, from a quantum system, while we care about what happens to the information *before measurement*. Note that the measurement axiom is the most philosophically "controversial," as it describes reality in a necessarily stochastic manner.

The Hilbert space setting allows us to use a variety of well-developed mathematical techniques without worrying too much about the physical realities underlying our system's phenomena. However, a Hilbert space has a fair amount of *structure*.[2] Brutalizing notation a bit, one could view the *algebraic* aspects of a Hilbert space $\mathcal{H}$ as the ordered, nested double

$$\underbrace{\left(\overbrace{\left(\underbrace{(\mathcal{H},+)}_{\text{group}},\mathbb{C}\right)}^{\text{vector space}},(\cdot,\cdot)\right)}_{\text{inner product space}},$$

ignoring the *topological* concerns. If none of this means anything to you, *do not worry*, especially considering no one writes out structures in this way! Still, we will use this nested deconstruction of the Hilbert space's algebraic structure to guide our discussion. As you can see, the algebraic structure of a Hilbert space is layered. The lowest level, an additive group, is built out of a set and an operation. We begin our study here. But first, we need to adopt a common language–that of informal set theory.

1.2. **Naive Set Theory.** Suppose we have a universe $\mathcal{U}$. We can think of a *set* $S$ as a collection of certain objects which live in $\mathcal{U}$. If an object is a *member* of $S$, we write $x \in S$, and if not, we write $x \notin S$, saying that $x$ is or is not an "element of" $S$, respectively. We will often define a set in the form

$$S := \{x \in \mathcal{U} : p(x)\},$$

saying "$S$ is all $x$ in $\mathcal{U}$ such that $p(x)$ holds." Now, given this loose definition of a set, there are a few operations we can perform to form new sets from old.

**Definition 1.5** (Set Arithmetic)**.** Let $A$ and $B$ be two arbitrary sets such that

$$A = \{x \in \mathcal{U} : p(x)\} \text{ and } B = \{x \in \mathcal{U} : q(x)\}.$$

(i) The set union of $A$ and $B$ is denoted $A \cup B$, and is defined as

$$A \cup B = \{x \in \mathcal{U} : p(x) \text{ or } q(x)\},$$

where "or" means either $A$, or $B$, *or both*.

(ii) The set intersection of $A$ and $B$ is denoted $A \cap B$, and is defined as

$$A \cap B = \{x \in \mathcal{U} : p(x) \text{ and } q(x)\}.$$

(iii) The set difference of $A$ minus $B$ is denoted $A \setminus B$, as is defined as

$$A \setminus B = \{x \in \mathcal{U} : p(x), \text{ but not } q(x)\}.$$

**Definition 1.6** (Subset)**.** Let $A$ be a set. We say that $B$ is a subset of $A$ ($B \subseteq A$) if every element in $B$ is also present in $A$.

**Definition 1.7** (Set Equality)**.** Let $A$ and $B$ be sets. If $A \subseteq B$ and $B \subseteq A$, then we say $A = B$.

[2]The precise notion of mathematical structure can be investigated in many ways. There are both organizational and foundational approaches. In the early twentieth century, a group of French mathematicians, under the pseudonym *Nicholas Bourbaki*, attempted to organize structure into three major components: algebraic, topological, and order. While this organizational framework become rather obsolete after the development of category theory, it can, at times, provide worthwhile intuition. Set theory and type theories form much of the foundational side, and in recent years, there have been efforts, such as homotopy type theory, which bridge the gap between foundational and organizational. Here, we avoid this deeper, philosophical study, instead using structure to guide us pedagogically.

Generally, to prove two sets are equal, we must show that $A \subseteq B$ and $B \subseteq A$. That is, we let $a \in A$ be an arbitrary element, and then show $a \in B$. Then, we let $b \in B$ be arbitrary, and show $b \in A$.

**Definition 1.8** (Empty Set)**.** The simplest set is the empty set $\{\ \}$, denoted $\varnothing$.

**Example 1.9** (Useful Sets)**.** You should be familiar with some standard sets:

(i) $\mathbb{Z}$ is the set of *integers* $\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$
(ii) $\mathbb{Q}$ is the set of *rational numbers*, or integer fractions,
$$\mathbb{Q} := \left\{\frac{a}{b} : a, b \in \mathbb{Z}\right\}.$$
(iii) $\mathbb{R}$ is the set of *real numbers*, including $\mathbb{Q}$ and all numbers "between" the rationals–the *irrationals*.
(iv) $\mathbb{C}$ is the set of *complex numbers*
$$\mathbb{C} := \{a + bi : a, b \in \mathbb{R} \text{ and } i^2 = 1\}$$

Using our definitions above, it is clear that $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

**Definition 1.10** (Disjoint)**.** If the intersection $A \cap B = \varnothing$, then we say $A$ and $B$ are disjoint. That is, there are no elements both in $A$ and in $B$.

**Definition 1.11** (Product)**.** The (Cartesian, direct) product of sets $A, B$ is the set
$$A \times B := \{(a, b) : a \in A \text{ and } b \in B\}$$
of ordered pairs.

**Example 1.12** (Cartesian Plane)**.** The plane $\mathbb{R}^2$ is precisely $\mathbb{R} \times \mathbb{R}$, all ordered pairs with real entries.

**Definition 1.13** (Function)**.** Let $A, B$ be sets. A function $f$ from $A$ to $B$ sends each element $a$ in $A$ to a single element $b$ in $B$ by some formula $f(a)$. We will write
$$f : A \to B \quad \text{or} \quad A \xrightarrow{f} B$$
for a function between $A$ and $B$.

To show the rule for $f$, we will often define a function using the notation
$$\begin{aligned} A &\xrightarrow{f} B \\ a &\longmapsto f(a), \end{aligned}$$
where $a \mapsto f(a)$ shows how an element $a$ moves "through the function $f$."

**Definition 1.14** (Image, Preimage)**.** Given a function $f : A \to B$, we define

(i) the image $f(A) := \{b \in B : f(a) = b \text{ for some } a \in A\}$.
(ii) for any subset $B' \subseteq B$, the preimage $f^{-1}(B') := \{a \in A : f(a) \in B'\}$.

**Definition 1.15** (Injective)**.** Let $f : A \to B$ be a function and $a, a' \in A$. Then, $f$ is injective if $f(a) = f(a')$ implies $a = a'$.

**Definition 1.16** (Surjective)**.** Let $f : A \to B$ be a function. If $f(A) = B$, then $f$ is surjective. That is, the image of $f$ somehow "fills" all of the codomain $B$.

**Definition 1.17** (Bijective)**.** A function $f : A \to B$ is bijective if it is injective and surjective.

*Remark* 1.18 (Arrow Notation). We will sometimes use decorated arrows for certain types of functions:[3]

(i) If $f$ is injective, we write $f : A \rightarrowtail B$, saying $f$ maps $A$ "into" $B$.
(ii) If $f$ is surjective, we write $f : A \twoheadrightarrow B$, saying $f$ maps $A$ "onto" $B$.
(iii) If $f$ is a bijection, we write $f : A \xrightarrow{\sim} B$, saying $f$ puts $A$ in "correspondence" with $B$.

**Definition 1.19** (Composition)**.** Given two functions $f : A \to B$ and $g : B \to C$, the composition $g \circ f$, read "$g$ after $f$," is a function $g \circ f : A \to C$ such that $a \mapsto g(f(a))$.

If you prefer to write functions pictorially, we can say that the composition $g \circ f$ is the function such that the following diagram commutes.[4]

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & B \arrow[r, "g"] & C \end{tikzcd}$$

**Example 1.20** (Identity Map)**.** For instance, the identity map $\mathrm{id}_A : A \xrightarrow{\sim} A$ sending $a \mapsto a$ is certainly a bijection between $A$ and itself.

**Definition 1.21** (Invertible)**.** If $f : A \xrightarrow{\sim} B$ is a bijection, then it is invertible with inverse $f^{-1} : B \xrightarrow{\sim} A$. The inverse $f^{-1}$ is the unique function such that $f^{-1} \circ f = \mathrm{id}_A$ and $f \circ f^{-1} = \mathrm{id}_B$.

1.3. **From Sets to Groups: Adding an Operation.** Now, given a set $S$, how can we go about performing operations on its members? For instance, if $S = \mathbb{Z}$, the integers, how should we define a familiar operation like addition? Well, addition takes two integers as inputs, and spits another one out as its output. If $a$ and $b$ are our inputs, and $a + b$ the output, then this action of addition can be written as $(a, b) \mapsto a + b$, so addition is a function

$$\begin{aligned} \mathbb{Z} \times \mathbb{Z} &\xrightarrow{(-)+(-)} \mathbb{Z} \\ (a, b) &\longmapsto a + b. \end{aligned}$$

In general, we can define a *binary operation* $\cdot$ on a set $S$ by taking $(-) \cdot (-) : S \times S \to S$, where $(s_1, s_2) \mapsto s_1 \cdot s_2$. Such a definition intrinsically requires *closure*, as if $g \notin S$, then we cannot have addition taking $s_1 + s_2 = g$ outside of the codomain $S$.

**Definition 1.22** (Group)**.** A group is a pair $(G, \cdot)$, where $G$ is a set and $(-) \cdot (-) : G \times G \to G$ is a binary operation, satisfying

(i) *associativity*: for all $g_1, g_2, g_3 \in G$, $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$.
(ii) *identity*: there exists a unique $e \in G$ so that for all $g \in G$, $e \cdot g = g \cdot e = g$.
(iii) *inverse*: for all $g \in G$, there exists an inverse $g^{-1}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Note that we used the symbol $\cdot$ for our binary operation, but really any symbol could have been used. The addition symbol $+$ is reserved for a certain type of group, which we will now introduce.

[3]If $A \subseteq B$, you will also see the arrow $\iota : A \hookrightarrow B$ to denote inclusions, where $\iota : a \mapsto a$. Since inclusions are closely related to injections, it is common practice to write $A \hookrightarrow B$ in place of $A \rightarrowtail B$, and vice-versa. These arrows tend to be more common in mathematics using lots of diagrams, such as in any category-heavy subject like algebraic topology. I will preface any usage of these arrows with the respective adjective.

[4]That is, if we trace an input either way along the diagram, you get the same output.

**Definition 1.23** (Abelian Group)**.** A group $(G, \cdot)$ is called abelian if for all $g_1, g_2 \in G$, $g_1 \cdot g_2 = g_2 \cdot g_1$. That is, the group $G$ is *commutative* under the operation $(-) \cdot (-) : G \times G \to G$.

**Definition 1.24** (Additive Group)**.** If we use the symbol $+$ for our operation and $(G, +)$ is an abelian group, then we call it additive.

*Remark* 1.25 (Group Notation). We will usually suppress the ordered pair notation $(G, +)$ for groups, instead just writing $G$. Technically, $G$ is just a set, but the operation is often apparent from context. Also, if $G$ uses $\cdot$ as its operation, we write $1_G$ for its identity $e$, whereas if it uses $+$ as its operation, we write $e = 0_G$. Similarly, multiplicative groups have inverses $g^{-1}$, while additive groups use $-g$. These notations stay consistent with how we think of addition and multiplication in $\mathbb{Z}$.

**Example 1.26** (Common Groups)**.**

(i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are all groups with $+$ as the operation.
(ii) $\mathbb{Z}/n\mathbb{Z} := \{0, 1, 2, \ldots, n-1\}$ is a group with addition modulo $n$ as the operation.
(iii) $\mathbb{C}^n := \{(z_1, \ldots, z_n) : z_i \in \mathbb{C}\}$ is a group under addition.
(iv) The set of square complex matrices

$$\mathbb{M}_n(\mathbb{C}) := \left\{ \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{pmatrix} : z_{ij} \in \mathbb{C} \right\}$$

is a group under addition.[5] One useful example is $\mathbb{M}_2(\mathbb{C})$, the $2 \times 2$ matrices.

**Definition 1.27** (Homomorphism)**.** A function $\varphi : (G_1, +_1) \to (G_2, +_2)$ between groups is called a homomorphism if it preserves the operations. That is, $\varphi(g +_1 g') = \varphi(g) +_2 \varphi(g')$.[6]

**Definition 1.28** (Isomorphism)**.** A homomorphism $\varphi : G_1 \to G_2$ between groups is an isomorphism if it has an inverse $\varphi^{-1}$ (or if it is bijective).

Given groups $G_1, G_2$, then the pair is *isomorphic* if there exists an isomorphism $\varphi : G_1 \xrightarrow{\sim} G_2$. In this case, we write $G_1 \simeq G_2$. Similarly, we could say that two sets $S_1, S_2$ are isomorphic if there exists a bijection $f : S_1 \xrightarrow{\sim} S_2$ between them. We will soon see one more structure in which a notion of isomorphism develops. In general, isomorphisms are functions which preserve the structure and have inverses which also preserve the structure.

**Definition 1.29** (Subgroup)**.** A subset $H \subseteq G$ of a group $G$ is called a subgroup if it is a group under the same operation. We denote this by $H \leq G$.

**Example 1.30** (Trivial Group)**.** Every group $G$ has a "trivial" subgroup $\{e\} \leq G$ consisting of just the identity element. Note that If $\{e'\} \leq H$ is a different trivial subgroup, we have an isomorphism $\varphi : \{e\} \xrightarrow{\sim} \{e'\}$ sending $e \mapsto e'$, meaning $\{e\} \simeq \{e'\}$. Thus, we can speak of *the trivial group*, which is isomorphically contained in all other groups.

Note that $\varnothing$ is *not* a group. This is because the group axioms insist on a group having an identity element, yet $\varnothing$ has *no elements*, by definition.

[5] It is not, however, a group under a multiplication, since matrices often are not invertible. Instead, we would have to consider the set of $n \times n$ invertible matrices $\mathrm{GL}_n(\mathbb{C}) \subseteq \mathbb{M}_n(\mathbb{C})$. It turns out, this *general linear group*, is, in fact, a group under matrix multiplication.

[6] The set of all group homomorphisms from $G_1$ to $G_2$ is written $\mathrm{Hom}_{\mathsf{Grp}}(G_1, G_2)$.

**Example 1.31** (Full Group)**.** Since $G \subseteq G$ is always a subset, if $G$ is a group, then $G \leq G$ is also a subgroup.

**Proposition 1.32** (Conditions for Subgroups)**.** *A subset $H \subseteq G$ is a subgroup if and only if it has the identity element $e$, is closed under addition, and is closed under taking inverses.*

**Definition 1.33** (Kernel)**.** Given a group homomorphism $\varphi : G_1 \to G_2$, we define the kernel of $\varphi$ to be the set

$$\ker \varphi := \{g \in G_1 : \varphi(g) = e_{G_2}\},$$

where $e_{G_2} \in G_2$ is the identity.

In a sense, the kernel of a homomorphism $\varphi : G_1 \to G_2$ is the set of elements in the domain $G_1$ which are "killed" by $\varphi$.

**Definition 1.34** (Center)**.** The center $\mathcal{Z}(G)$ of a group $G$ is defined as

$$\mathcal{Z}(G) := \{g \in G : gx = xg \text{ for all } x \in G\}.$$

**Definition 1.35** (Centralizer)**.** The centralizer $C_G(S)$ of a set $S \subseteq G$ is

$$C_G(S) := \{g \in G : gs = sg \text{ for all } s \in S\}.$$

It should be clear that the centralizer $C_G(G) = \mathcal{Z}(G)$, the center.

**Definition 1.36** (Normalizer)**.** The normalizer $\mathcal{N}_G(S)$ of a subset $S \subseteq G$ is

$$\mathcal{N}_G(S) := \{g \in G : gSg^{-1} = S\},$$

where

$$gSg := \{gsg^{-1} : s \in S\}.$$

**Proposition 1.37** (Useful Subgroups)**.** *The kernel, center, centralizer, and normalizer are all subgroups.*

**Definition 1.38** (Subgroup Generated by Subset)**.** Let $S \subseteq G$ be a subset of a group. Then, the subgroup generated by $S$ in $G$, denoted $\langle S \rangle$, is the smallest subgroup of $G$ containing $S$ as a subset. That is,

$$\langle S \rangle := \bigcap_{S \subseteq H \leq G} H,$$

the intersection of all such subgroups.

If $S = \{g_1, \ldots, g_n\} \subseteq G$ is a finite subset, then we will write

$$\langle g_1, \ldots, g_n \rangle \leq G$$

for the subgroup generated by $S$.

*Remark* 1.39 (Alternative Characterization). Note that the subgroup generated by $S \subseteq G$ is the subgroup of all elements in $G$ which can be expressed as finite products of elements in $S$ and their inverses.

**Example 1.40** (Pauli Group)**.** We end our discussion of groups by defining the "most important" group when doing error correction. Consider the $2 \times 2$ complex matrices, called the *Pauli matrices*,

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{M}_2(\mathbb{C}).$$

Define the group generated by these matrices $\mathcal{P} := \langle X, Y, Z \rangle$, using matrix multiplication as the operation. Then, you may check that $\mathcal{P} = \{\pm I, \pm X, \pm Y, \pm Z, \pm iI, \pm iX, \pm iY, \pm iZ\}$, where

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is the $2 \times 2$ identity matrix. We call $\mathcal{P}$ the *Pauli group*.

1.4. **From Groups to Vector Spaces: Adding Complex Linearity.** Finally, we make the last leaps from sets to the finite dimensional Hilbert spaces of our axioms.

**Definition 1.41** (Vector Space)**.** A vector space over $\mathbb{C}$ (or a $\mathbb{C}$-linear space) is a pair $((\mathcal{H}, +), \cdot)$, where $(\mathcal{H}, +)$ is an additive group and $(-) \cdot (-) : \mathbb{C} \times \mathcal{H} \to \mathcal{H}$ is an "action" by the *field* of complex numbers, written with juxtaposition,[7] satisfying

(i) *compatibility*: for all $v \in \mathcal{H}$ and $c_1, c_2 \in \mathbb{C}$, $c_1(c_2 v) = (c_1 c_2)v$.
(ii) *identity*: for all $v \in \mathcal{H}$, $1v = v$.
(iii) *distributivity*: for all $v_1, v_2 \in \mathcal{H}$ and $c \in \mathbb{C}$, $c(v_1 + v_2) = cv_1 + cv_2$.
(iv) *distributivity, again*: for all $v \in \mathcal{H}$ and $c_1, c_2 \in \mathbb{C}$, $(c_1 + c_2)v = c_1 v + c_2 v$.

*Remark* 1.42 (Vector Terminology). We call the elements $v \in \mathcal{H}$ *vectors*, the elements $c \in \mathbb{C}$ *scalars*, and any element $g \in \mathcal{H}$ of the form

$$g = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = \sum_{i=1}^{n} c_i v_i,$$

where $v_i \in \mathcal{H}$ and $c_i \in \mathbb{C}$, a *linear combination*.

**Definition 1.43** (Linear Independence)**.** A subset $S \subseteq \mathcal{H}$ of a vector space is linearly independent if there does not exist an $s \in S$ so that $s$ is a linear combination of elements in $\mathcal{H}$.

When we looked at sets and groups, there was an idea of a sub-structure. As you might expect, the same notion exists for vector spaces.

**Definition 1.44** (Subspace)**.** A subset $\mathcal{W} \subseteq \mathcal{H}$ is a subspace if it is closed under addition in $\mathcal{H}$ and multiplication from $\mathbb{C}$.[8]

*Remark* 1.45 (Subspace Notation). Unlike groups, but like other objects such as rings and fields, we do not have a special notation for being a subspace of a vector space.

**Definition 1.46** (Span)**.** Let $S \subseteq \mathcal{H}$, as before. Then, we define the span of $S$ to be the subspace

$$\operatorname{span} S := \left\{ \sum_{i=1}^{n} c_i s_i : s_i \in \mathcal{H} \text{ and } c_i \in \mathbb{C} \right\}.$$

**Proposition 1.47** (Equivalence of Span)**.** *The span of a subset $S \subseteq \mathcal{H}$ is*

$$\operatorname{span} S = \bigcap_{S \subseteq \mathcal{W} \subseteq \mathcal{H}} \mathcal{W},$$

*the smallest subspace of $\mathcal{H}$ containing $S$.*[9]

[7] In general, we can put any field here. Some difficulties arise when talking about finite fields, especially those of characteristic 2, but we omit this generality purely because they are irrelevant to our discussion.

[8] This definition forces the fact that if $\mathcal{W}_1, \mathcal{W}_2 \subseteq \mathcal{H}$ are subspaces, then $\mathcal{W}_1 \cap \mathcal{W}_2 \subseteq \mathcal{H}$ is too.

[9] Taking the span of a subset is similar to our notion of generating a subgroup. In this light, it is intuitive to write $\operatorname{span}\{s_1, \ldots, s_n\}$ for the span of a finite subset $\{s_1, \ldots, s_n\} \subseteq \mathcal{H}$.

**Definition 1.48** (Basis)**.** A basis $\beta$ of a vector space $\mathcal{H}$ is a linearly independent, minimal spanning set of $\mathcal{H}$, where minimality is with respect to cardinality.

**Definition 1.49** (Dimension)**.** The dimension of a vector space $\mathcal{H}$ is the cardinality, or size, of any basis $\beta$ of $\mathcal{H}$. If the size $|\beta| = n$, then $\dim \mathcal{H} := n$.

**Theorem 1.50** (Basis Existence)**.** *Every vector space has a basis.*

**Definition 1.51** (Linear Transformation)**.** Given two $\mathbb{C}$-linear spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, a function $T : \mathcal{H}_1 \to \mathcal{H}_2$ is called a linear transformation (or vector space homomorphism) if it satisfies

$$T(v_1 + cv_2) = T(v_1) + cT(v_2)$$

for all $v_1, v_2 \in \mathcal{H}_1$ and $c \in \mathbb{C}$.

**Definition 1.52** (Linear Operator)**.** A linear transformation $T : \mathcal{H} \to \mathcal{H}$ between a space and itself is called a linear operator.[10]

**Definition 1.53** (Kernel)**.** The kernel (or null space) of a linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$, denoted $\ker T$, is defined as the $\mathcal{H}_1$ fiber over 0:

$$\ker T := \{v \in \mathcal{H}_1 : Tv = 0\}.$$

**Definition 1.54** (Image)**.** The image, of a linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$ is

$$T(\mathcal{H}_1) := \{w \in \mathcal{H}_2 : Tv = w \text{ for some } v \in \mathcal{H}_1\}.$$

**Definition 1.55** (Isomorphism)**.** A linear map $T : \mathcal{H}_1 \xrightarrow{\sim} \mathcal{H}_2$ is called an isomorphism if it has an inverse $T^{-1}$ (or is bijective).

If an isomorphism $T : \mathcal{H}_1 \xrightarrow{\sim} \mathcal{H}_2$ exists, then we say the pair is *isomorphic* and write $\mathcal{H}_1 \simeq \mathcal{H}_2$. Now that we have a notion of "sameness" for our structure, as we did with sets and groups, we define the *product* of spaces $\mathcal{W}_1, \mathcal{W}_2$ to be

$$\mathcal{W}_1 \times \mathcal{W}_2 := \{(w_1, w_2) : w_i \in \mathcal{W}_i\},$$

all ordered pairs of elements in $\mathcal{W}_i$, respectively. For a large collection of spaces $\{\mathcal{W}_i\}_{i \in I}$, the space

$$\prod_{i \in I} \mathcal{W}_i = \mathcal{W}_1 \times \mathcal{W}_2 \times \cdots$$

is the product of all the $\mathcal{W}_i$, where the operations are taken component-wise, as you would expect.

**Definition 1.56** (Direct Sum)**.** Let $\{\mathcal{W}_i\}_{i \in I}$ be an indexed family of vector spaces. Then, we define the direct sum (or coproduct) of these spaces to be the space

$$\bigoplus_{i \in I} \mathcal{W}_i := \{(w_1, w_2, \dots) : \text{all but finitely many } w_i = 0\}.$$

Since both $\bigoplus \mathcal{W}_i$ and $\prod \mathcal{W}_i$ are spaces of tuples, when the indexing set $I = \{1, \dots, n\}$ is finite, the spaces are isomorphic $\bigoplus \mathcal{W}_i \simeq \prod \mathcal{W}_i$. Then,

$$\mathbb{C}^n \simeq \mathbb{C}^{\oplus n} = \mathbb{C} \oplus \mathbb{C} \underbrace{\oplus \cdots \oplus}_{n \text{ times}} \mathbb{C}.$$

**Proposition 1.57** (Sum of Dimension)**.** *The dimension* $\dim(\mathcal{H}_1 \oplus \mathcal{H}_2) = \dim(\mathcal{H}_1) + \dim(\mathcal{H}_2)$.

[10] This is one way we use the word operator. Sometimes, we will call any map between inner product spaces an operator.

*Remark* 1.58 (Note on Dimension). Let $\{\mathbb{C}\}_{i\in I} \subseteq \mathcal{H}$ be a subfamily of all the one dimensional subspaces of $\mathcal{H}$, up to isomorphism Take a direct sum decomposition of a space

$$\mathcal{H} = \bigoplus_{i\in I} \mathbb{C}.$$

Since $\mathbb{C}$ is the *base field* of the vector space $\mathcal{H}$, which has $\dim \mathbb{C} = 1$ over itself, we, in a sense, cannot break the direct sum up any further. Thus, taking dimensions, we see that

$$\dim \mathcal{H} = \sum_{i\in I} \dim \mathbb{C} = \sum_{i\in I} 1 = |I|.$$

In the finite case, this means

$$\dim \mathcal{H} = \sum_{i=1}^{n} 1 = 1 + 1 \underbrace{+ \cdots +}_{n \text{ times}} 1 = n.$$

Thus, the dimension of a vector space is *exactly* how many copies of the field $\mathbb{C}$ can be fit into a direct sum decomposition. If this number is infinite, then so is $\dim \mathcal{H}$.

**Example 1.59** ($\mathbb{C}^n$ Dimension)**.** Consider the space $\mathbb{C}^n$. We know, from experience with bases, that $\dim \mathbb{C}^n = n$. Yet, we could also have taken a direct sum decomposition

$$\mathbb{C}^n \simeq \mathbb{C}^{\oplus n},$$

which is a direct sum decomposition with $n$ copies of $\mathbb{C}$. This is a nice sanity check that our reinterpretation of dimension makes sense.

**Theorem 1.60** (Finite Spaces)**.** *Every space of dimension* $\dim \mathcal{H} = n < \infty$ *is isomorphic to* $\mathbb{C}^n$.

**Proposition 1.61** (Tautological Sum Isomorphism)**.** *Given a direct sum decomposition*

$$\mathcal{H} = \bigoplus_{i=1}^{n} \mathcal{H}_i,$$

*every element* $v \in \mathcal{H}$ *can be written in the form*

$$v = w_1 + w_2 + \cdots + w_n,$$

*where* $w_i \in \mathcal{H}_i$.

We now have a way to translate between the tuples of the direct sum characterization and the sums of elements of the direct summands.

**Definition 1.62** (Projection)**.** Given a vector space decomposition $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_n$, the projection (or projector) onto $\mathcal{H}_1$ is given by the transformation

$$\begin{aligned} \bigoplus_{i=1}^{n} \mathcal{H}_i &\xrightarrow{P_1} \mathcal{H}_1 \\ \sum_{i=1}^{n} v_i &\longmapsto v_1. \end{aligned}$$

Intuitively, projections are a method of lossy compression, ignoring all of the information from all subspaces in the coproduct *except* the space we are projecting onto. We now focus in on finite dimensional spaces, as those are the primary spaces of interest in quantum error correction. Hereafter, unless stated more generally, assume all spaces are finite dimensional.

**Proposition 1.63** (Matrix Representation)**.** *Every linear transformation between $\mathcal{H}_1$ and $\mathcal{H}_2$, having dimensions $n$ and $m$, respectively, can be realized as a matrix in $\mathbb{M}_{m\times n}(\mathbb{C})$. If $\mathcal{H}_1 \simeq \mathcal{H}_2$, then*

$$\mathbb{C}^n \simeq \mathcal{H}_1 \simeq \mathcal{H}_2 \simeq \mathbb{C}^m,$$

*so $n = m$ and our linear operators $T : \mathcal{H}_1 \to \mathcal{H}_1$ can be represented by matrices in $\mathbb{M}_n(\mathbb{C})$. The converse is also true, so the linear maps and matrices are in bijective correspondence.*

*Remark* 1.64. A common way to depict the matrix representation of $T : \mathcal{H}_1 \to \mathcal{H}_2$ is $[T]_\beta^\gamma$, where $\beta, \gamma$ are bases for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then, assuming $\dim \mathcal{H}_1 = n$ and $\dim \mathcal{H}_2 = m$, we have that the following diagram commutes:

$$\begin{tikzcd}
\mathcal{H}_1 \arrow[r, "T"] \arrow[d, "\simeq"'] & \mathcal{H}_2 \arrow[d, "\simeq"] \\
\mathbb{C}^n \arrow[r, "{[T]_\beta^\gamma}"'] & \mathbb{C}^m
\end{tikzcd}$$

where the isomorphisms are precisely the action of picking the bases $\beta, \gamma$. Yet, there is no recipe for picking such an isomorphism, as we have many choices of a basis. We tend to say "the action of picking a basis is not *canonical*."

**Definition 1.65** (Space of Linear Maps)**.** The space of linear maps $T : \mathcal{H}_1 \to \mathcal{H}_2$ is denoted $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$.[11]

**Proposition 1.66** (Linear Maps Form a Vector Space)**.** *As indicated by naming, the space $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is a vector space.*

Using the previous proposition, we see that if $\dim \mathcal{H} = n$, then $\mathcal{L}(\mathcal{H}) \simeq \mathbb{M}_n(\mathbb{C})$. Now, we need to capture a notion of being orthogonal in our vector space $\mathcal{H}$.

**Definition 1.67** (Inner Product)**.** Take arbitrary $v_1, v_2, v_3 \in \mathcal{H}$ and $c_1, c_2 \in \mathbb{C}$. An inner product on the $\mathbb{C}$-linear space $\mathcal{H}$ is a map

$$(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$

satisfying

(i) *conjugate symmetry*: $(v_1, v_2) = (v_2, v_1)^*$.[12]
(ii) *linearity in the second argument*:

$$(v_3, c_1 v_1 + c_2 v_2) = c_1(v_3, v_1) + c_2(v_3, v_2).$$

(iii) *positive semi-definiteness*: $(v_1, v_1) \geq 0$, where equality holds if and only if $v_1 = 0$.

*Remark* 1.68 (Dot Product). The inner product of arbitrary vector spaces captures the idea of the dot product in $\mathbb{R}^n$. In particular, like with the dot product, two vectors $v_1, v_2 \in \mathcal{H}$ are said to be *orthogonal* if $(v_1, v_2) = 0$ (zero inner product).

[11] In the literature, you may see the notation $\mathrm{Hom}_{\mathbb{C}}(\mathcal{H}_1, \mathcal{H}_2)$, instead. This is more common in algebra, whereas the notation we introduced is more common in information theory. Later, we will see the notation used in functional analysis, which we will adopt.

[12] Here, the star is the complex conjugate, taking $(a + bi)^* = a - bi$. In mathematics notation, this is denoted $\overline{a + bi}$. I do not know why the difference exists. I will use physics notation here, solely because it matches much of the literature. Then, for *adjoints*, I will use $\dagger$ instead of $*$, staying consistent. I do avoid using Dirac notation, however, just out of preference.

**Definition 1.69** (Inner Product Space)**.** An inner product space is a pair $(\mathcal{H}, (\cdot,\cdot))$ consisting of a vector space $\mathcal{H}$ and an inner product $(\cdot,\cdot)$.

Using the inner product, we can actually extract information about *sizes* of vectors in our space. For a vector $v \in \mathcal{H}$, an inner product space, define $\|v\| := \sqrt{(v,v)}$. This is called the *norm* induced by the inner product.

**Definition 1.70** (Hilbert Space)**.** A (complex) Hilbert space $\mathcal{H}$ is an (complex) inner product space which is complete with respect to the (metric from the) norm induced by the inner product.

**Theorem 1.71** (Finite Dimensional Hilbert Space)**.** *If $\mathcal{H}$ is a finite dimensional inner product space, it is always complete with respect to the induced norm.*

**Definition 1.72** (Bounded Operators)**.** In a finite dimensional Hilbert space $\mathcal{H}$ isomorphic to $\mathbb{C}^n$, the set of bounded operators $\mathbb{B}(\mathcal{H})$ is *exactly* the set of linear operators $\mathcal{L}(\mathcal{H})$.

*Remark* 1.73 (Boundedness). In the general theory of Hilbert spaces, where dimension can be non-finite, we must define what it means for linear operators to be bounded. However, in finite dimensions, all operators are bounded, so we do not need more theory.

*Remark* 1.74 (C*-algebras). Since $\mathcal{L}(\mathbb{C}^n) \simeq \mathbb{M}_n(\mathbb{C})$, we have that $\mathbb{B}(\mathbb{C}^n) \simeq \mathbb{M}_n(\mathbb{C})$. In the literature, you may come across the notion of a C*-algebra. Firstly, an algebra is a vector space along with another operation denoted by multiplication. Every finite dimensional C*-algebra is isomorphic to a subalgebra of $\mathbb{M}_n(\mathbb{C})$. If you come across the terminology, you want to think of C*-algebras as subalgebras of $\mathbb{B}(\mathcal{H})$, and in turn, think of subalgebras of $\mathbb{M}_n(\mathbb{C})$.

**Definition 1.75** (Adjoint)**.** Let $\mathcal{H}$ be a Hilbert space. Consider an operator $T \in \mathbb{B}(\mathcal{H})$. The (Hermitian) adjoint of $T$ is an operator $T^\dagger \in \mathbb{B}(\mathcal{H})$ is defined to satisfy the relation

$$(Tv_1, v_2) = (v_1, T^\dagger v_2).$$

*Remark* 1.76 (Conjugate Transpose). In finite dimensions, taking the adjoint $T^\dagger$ yields the same thing as the conjugate transpose $(T^*)^t$:

$$T^\dagger = \begin{pmatrix} z_{11}^* & z_{12}^* & \cdots & z_{1n}^* \\ z_{21}^* & z_{22}^* & \cdots & z_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1}^* & z_{n2}^* & \cdots & z_{nn}^* \end{pmatrix}^t = \begin{pmatrix} z_{11}^* & z_{21}^* & \cdots & z_{n1}^* \\ z_{12}^* & z_{22}^* & \cdots & z_{n2}^* \\ \vdots & \vdots & \ddots & \vdots \\ z_{1n}^* & z_{2n}^* & \cdots & z_{nn}^* \end{pmatrix}.$$

**Definition 1.77** (Dual Space)**.** Given a vector space $\mathcal{H}$, the dual space

$$\mathcal{H}^* := \{v^\dagger : \mathcal{H} \to \mathbb{C}\} = \mathcal{L}(\mathcal{H}, \mathbb{C})$$

is the space of *linear functionals*. If $\mathcal{H} \simeq \mathbb{C}^n$, an element $v^\dagger \in \mathcal{H}^*$ is precisely the *conjugate transpose* of $v \in \mathcal{H}$, turning the column vector to a row vector and taking conjugates.

*Remark* 1.78 (Working with Hilbert Spaces). Because our Hilbert spaces $\mathcal{H}$ are finite dimensional when modeling finite quantum systems, we can reduce our study to just inner product spaces! Notably, it is reasonable to mentally work using the isomorphism $\mathcal{H} \simeq \mathbb{C}^n$. The inner product is given by

$$(v_1, v_2) = v_1^\dagger v_2 = \sum_{i=1}^{n} v_{1_i}^* v_{2_i}.$$

We have now reviewed the basics from sets, to additive groups, to finite dimensional vector spaces, to inner product spaces (finite dimensional Hilbert spaces). Recalling our structure listing from before, we have made it fully up the ladder. The last order of business is to build the Hilbert space language of operators up to fully parse the four axioms of quantum theory, stated in the introduction.

**Definition 1.79** (Self-Adjoint)**.** An operator $T \in \mathbb{B}(\mathcal{H})$ is called self-adjoint (or Hermitian) if $T^\dagger = T$.

**Definition 1.80** (Eigen-)**.** Let $T \in \mathbb{B}(\mathcal{H})$ be an operator on a finite dimensional Hilbert space. Suppose $Tv = \lambda v$ for some $v \in \mathcal{H}$ and $\lambda \in \mathbb{C}$. Then, $v$ is an eigenvector and $\lambda$ is an eigenvalue.

**Theorem 1.81** (Spectral Theorem)**.** *Let $\mathcal{H}$ be a finite dimensional Hilbert space. Let $T \in \mathbb{B}(\mathcal{H})$ be self-adjoint. Then, there exits an orthonormal* (*orthogonal with unit-norm*) *basis for $\mathcal{H}$ of eigenvectors $v_i$ of $T$. Additionally, all the corresponding eigenvalues $\lambda_i \in \mathbb{R}$ are real.*

**Definition 1.82** (Isometry)**.** An operator $V \in \mathbb{B}(\mathcal{H})$ on a Hilbert space is an isometry if $V^\dagger V = I$, the identity on $\mathcal{H}$.

**Definition 1.83** (Unitary)**.** An operator is unitary if $U^\dagger U = UU^\dagger = I$.

**Proposition 1.84** (Equivalent Condition for Unitaries)**.** *An operator $U \in \mathbb{B}(\mathcal{H})$ is unitary if and only if $U : \mathcal{H} \twoheadrightarrow \mathcal{H}$ is an isometry.*

*Remark* 1.85 (Operator Isomorphism)*.* In the context of Hilbert spaces, our isomorphisms must not only preserve the vector space structure, but also the inner product. Unitary operators are precisely linear isomorphisms preserving the inner product, so they act as isomorphisms.

**Definition 1.86** (Positive Semi-Definite)**.** An operator $T \in \mathbb{B}(\mathcal{H})$ is positive semi-definite (or positive, denoted $T \geq 0$) if $(Tv, v) \geq 0$ for all $v \in \mathcal{H}$.

**Example 1.87** (Identity is Positive)**.** Consider the identity $I \in \mathbb{B}(\mathcal{H})$. Well, for all $v \in \mathcal{H}$, $(Iv, v) = (v, v) \geq 0$, so the identity is positive semi-definite by the definition of inner product.

Suppose we have a direct sum decomposition

$$\mathcal{H} \simeq \bigoplus_{i=1}^{n} \mathcal{H}_i.$$

If each of the $\mathcal{H}_i$ is such that $(v_1, v_2) = (v_2, v_1) = 0$ for all $v_i \in \mathcal{H}_i$, then we call the spaces *orthogonal* and dub the decomposition an *orthogonal decomposition*. Having an orthogonal decomposition is equivalent to having a collection of *orthogonal projectors* $\{P_i\}_{i=1}^n$ so that $P_i P_j = \delta_{ij} P_i$, where

$$\delta_{ij} := \begin{cases} 0, & i = j \\ 1, & i \neq j \end{cases}$$

is the *Kronecker delta*. Note that for projectors of an arbitrary direct sum, since $P_1(v_1 + \cdots + v_n) = v_1$, where $v_i \in \mathcal{H}_i$, we have that $P_1^2 = P_1$. That is, projectors are *idempotent*.

**Definition 1.88** (Trace)**.** Let $T \in \mathbb{B}(\mathcal{H})$ be a linear operator on a finite dimensional space. Then, taking the matrix representation, the trace $\operatorname{tr} T$ is the sum of the diagonal elements of the matrix.

**Proposition 1.89** (Trace and Eigenvalues)**.** *By simple linear algebra, this means the trace of an operator is precisely the sum of the eigenvalues.*

**Definition 1.90** (Density Operator)**.** A quantum density operator $\rho \in \mathbb{B}(\mathcal{H})$ is an operator such that $\operatorname{tr} \rho = 1$ and $\rho \geq 0$.

**Axiom 1.91** (State Space)**.** Any quantum system $Q$ is represented by a complex Hilbert space $\mathcal{H}^Q$, called the state space. States of the system are represented by unit-trace, positive semi-definite operators acting on $\mathcal{H}$, called density operators.

We tend to denote the subset of density operators $\mathcal{D}(\mathcal{H}) \subseteq \mathbb{B}(\mathcal{H})$. Similarly, the set of unitary operators is sometimes denoted $\mathcal{U}(\mathcal{H})$. Finally, we state two useful theorems we may use in later proofs.

**Theorem 1.92** (Singular Value Decomposition)**.** *Any operator $T \in \mathbb{B}(\mathcal{H})$ can be written as $T = U \Sigma V^\dagger$, where $U, V^\dagger \in \mathcal{U}(\mathcal{H})$ and $\Sigma$ is diagonal. The diagonal entries denoted $\sigma_i := \sqrt{\lambda_i}$ are called singular values, and are precisely the nonzero eignvalues of $T$.*

**Theorem 1.93** (Polar Decomposition)**.** *Any operator $T \in \mathbb{B}(\mathcal{H})$ can be expressed in the form $T = U\sqrt{T^\dagger T}$, where $U \in \mathcal{U}(\mathcal{H})$ is a unitary operator.*

## 2. The Subspace Condition

Over the course of the preliminaries, we have seen all of the mathematics needed to understand the axioms of quantum theory, except the second. We are missing tensor products. Once we can form these *composite quantum systems*, we will be ready to study error channels.

### 2.1. Tensor Products and Quantum Channels.

**Definition 2.1** (Bilinear)**.** A function $B : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$ between vector spaces over $\mathbb{C}$ is bilinear if for all $w \in \mathcal{H}_2$, the function $v \mapsto B(v, w)$, and for all $v \in \mathcal{H}_1$, the function $w \mapsto B(v, w)$ are linear transformations.

**Definition 2.2** (Universal Property of Tensor Product)**.** Let $\mathcal{H}_1, \mathcal{H}_2$ be two vector spaces. Define a new, tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ with a bilinear function

$$\begin{aligned} \mathcal{H}_1 \times \mathcal{H}_2 &\xrightarrow{(-)\otimes(-)} \mathcal{H}_1 \otimes \mathcal{H}_2 \\ (v, w) &\longmapsto v \otimes w \end{aligned}$$

such that for all bilinear maps $h : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$, there exists a *unique* linear map $\overline{h} : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_3$ such that $h = \overline{h} \circ \otimes$.

$$\begin{tikzcd} \mathcal{H}_1 \times \mathcal{H}_2 \arrow[r, "(-)\otimes(-)"] \arrow[dr, "h"'] & \mathcal{H}_1 \otimes \mathcal{H}_2 \arrow[d, dashed, "\exists!\ \overline{h}"] \\ & \mathcal{H}_3 \end{tikzcd}$$

Figure 2. Universal property of tensor product in diagram form. The tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the unique, up to isomorphism, space making the diagram commute

In practice, we will work with a more grounded interpretation of the tensor product. Let $\beta_1$ be a basis for $\mathcal{H}_1$ and $\beta_2$ be a basis for $\mathcal{H}_2$. Then, the set

$$\{v \otimes w : v \in \beta_1 \text{ and } w \in \beta_2\}$$

is a basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$ over $\mathbb{C}$. Hence, we "define" our tensor product space as[13]

$$\mathcal{H}^{AB} := \operatorname{span}\{v \otimes w : v \in \beta_1 \text{ and } w \in \beta_2\}.$$

Functionally, we will use the bilinearity of $\otimes$ to perform our computations.

**Axiom 2.3** (Multiple System)**.** Any pair of join quantum systems $A, B$ can be represented by a tensor product Hilbert space

$$\mathcal{H}^{AB} := \mathcal{H}^A \otimes \mathcal{H}^B.$$

If $\dim \mathcal{H}^A = n$ and $\dim \mathcal{H}^B = m$, then

$$\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B \simeq \mathbb{C}^n \otimes \mathbb{C}^m \simeq \mathbb{C}^{nm}.$$

**Definition 2.4** (Bipartite System)**.** We call a tensor system $\mathcal{H} := \mathcal{H}^A \otimes \mathcal{H}^B$ a bipartite system, as it represents two systems in composite.

**Definition 2.5** (Partial Trace)**.** Let $\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B$ be a bipartite system with an orthonormal basis $\{v_i \otimes w_i\}$. Then, the partial trace over system $A$ is the linear operator $\operatorname{tr}_A := \operatorname{tr} \otimes I^B$, taking in operators from $\mathbb{B}(\mathcal{H}^{AB})$ and outputting operators in $\mathbb{B}(\mathcal{H}^B)$.

When $\rho = a_1 a_2^\dagger \otimes b_1 b_2^\dagger$, then the partial trace over system $A$ yields

$$\operatorname{tr}_A(a_1 a_2^\dagger \otimes b_1 b_2^\dagger) = \operatorname{tr}\Big(a_1 a_2^\dagger\Big) b_1 b_2^\dagger,$$

"tracing out" $A$ and leaving $B$. More generally, if $\rho^{AB} \in \mathbb{B}(\mathcal{H}^{AB})$, then we define the reduced operator

$$\rho^B := \operatorname{tr}_A(\rho^{AB}) = \left(\sum_i v_i^\dagger \otimes I^B\right) \rho^{AB} \left(\sum_i v_i \otimes I^B\right) \in \mathbb{B}(\mathcal{H}^B),$$

where $I^B$ is the identity on $\mathcal{H}^B$. Similarly, the partial trace over system $B$ yields

$$\operatorname{tr}_B(a_1 a_2^\dagger \otimes b_1 b_2^\dagger) = b_1 b_2^\dagger \operatorname{tr}\Big(a_1 a_2^\dagger\Big)$$

for simple tensors,[14] and a reduced operator

$$\rho^A := \operatorname{tr}_B(\rho^{AB}) = \left(\sum_i I^A \otimes w_i^\dagger\right) \rho^{AB} \left(\sum_i I^A \otimes w_i\right) \in \mathbb{B}(\mathcal{H}^A)$$

for arbitrary $\rho^{AB} \in \mathbb{B}(\mathcal{H}^{AB})$.

*Remark* 2.6 (Marginal States). If $\rho^{AB} \in \mathcal{D}(\mathcal{H}^{AB})$ is a bipartite density operator, then

$$\rho^A := \operatorname{tr}_B(\rho^{AB}) \in \mathcal{D}(\mathcal{H}^A)$$

and

$$\rho^B := \operatorname{tr}_A(\rho^{AB}) \in \mathcal{D}(\mathcal{H}^B)$$

are called the *marginal sates* of the respective systems. If you are familiar with the terminology of marginal probabilities, the definitions are analogous.

[13]The construction using span can be made more formal. There are many constructions of tensor products of vectors spaces, but they tend to not be very illuminating for using them in computation.

[14]That is, one which can be written in the form $u \otimes v$. Note that the operator $v_1 v_2^\dagger$ is called an *outer product*, since it switches the order of the dagger from an inner product. You can imagine multiplying a column vector by a row vector on the right, which results in a matrix rather than a scalar.

*Remark* 2.7 (Tracing out the Environment). Intuition for the partial trace can come from thinking practically about quantum systems. Say we are performing a quantum experiment with system $A$. Well, the evolution axiom tells us that closed systems evolve unitarily. What if our system is not closed (as a real, physical system would not be)? That is, what if we have physical interactions coming in from the surrounding lab equipment, dust particles, and such (Fig. 3)? We can then *tensor together* our system $A$ and the environment system $E$. Our composite system $AE$ has Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^E$. Putting in a state $\rho^{AB} \in \mathcal{D}(\mathcal{H}^{AE})$, we can then consider this system to be closed, working as if it was. When it is time to measure, we trace out system $E$, only leaving us with the marginal state $\rho^A$

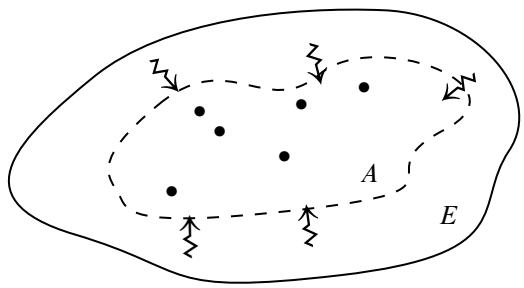


FIGURE 3. Experiment in system $A$, with interactions from environment system $E$. Forming the composite system $AE$ would "embed" $A$ into a closed system, ready for unitary evolution. The environment can be traced out via $\mathrm{tr}_E$.

**Definition 2.8** (Superoperator)**.** Since $\mathbb{B}(\mathcal{H})$ is a Hilbert space in its own right, we can consider bounded linear maps $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$. Such a map between operators is called a superoperator.

**Definition 2.9** (Trace Preserving)**.** Let $\Phi$ be a superoperator in $\mathbb{B}(\mathbb{B}(\mathcal{H}^A) : \mathbb{B}(\mathcal{H}^B))$. Then, if

$$\mathrm{tr}(\Phi(\rho)) = \mathrm{tr}(\rho),$$

for all $\rho \in \mathbb{B}(\mathcal{H}^A)$, we say $\Phi$ is trace preserving (TP).

*Remark* 2.10 (Positive Operator). Recall that an operator $\rho \in \mathbb{B}(\mathcal{H})$ is called *positive semi-definite* if $(Tv, v) \geq 0$ for all $v \in \mathcal{H}$. We write $\rho \geq 0$. It is a pervasive standard to simply call $\rho$ *positive*.

**Definition 2.11** (Positive Superoperator)**.** A superoperator $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ is called positive if $\Phi(\rho) \geq 0$ for all $\rho \geq 0$.

**Definition 2.12** ($k$-Positive Superoperator)**.** A superoperator $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ is called $k$-positive if the superoperator $\Phi \otimes I^{Q_k}$ is positive, where $I^{Q_k}$ is the identity on a space $\mathcal{H}^{Q_k} \simeq \mathbb{C}^k$.

**Definition 2.13** (Completely Positive Superoperator)**.** A superoperator is called completely positive (CP) if it is $k$-positive for all $k \in \{1, 2, \dots\}$.

**Definition 2.14** (Quantum Channel)**.** A superoperator is a quantum channel if it is completely positive and trace preserving (CPTP).

*Remark* 2.15 (Why Channels?). The question of why we use channels can be answered in many ways. There is a physical "naturality" to the process of channels. One good answer, is that they send our density matrices to density matrices, no matter how we couple an environment to them. Recall that a density matrix $\rho \in \mathcal{D}(\mathcal{H})$ is of unit-trace and positive. Since a channel $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})$ is trace preserving, $\Phi(\rho)$ is also of unit-trace. Similarly, if we couple an environment Hilbert space $\mathcal{H}^E$ to our experimental system Hilbert space $\mathcal{H}^A$, we can run the state $\rho^{AE}$ through the channel $\Phi : \mathbb{B}(\mathcal{H}^{AE}) \to \mathbb{B}(\mathcal{H}^{AE})$. Not only will the state be unit-trace, it will still be positive after tracing out $E$, as channels are completely positive. Thus, $\mathrm{tr}_E\, \Phi(\rho^{AE}) \in \mathcal{D}(\mathcal{H}^A)$. From a probabilistic perspective, we can interpret $0 \leq \mathrm{tr}(\Phi(\rho)) \leq 1$ to give us the probability of the process of the channel $\Phi$ occurring.

**Example 2.16** (Unitary Operators)**.** Unitary operators are channels.

**Example 2.17** (Partial Trace)**.** The partial trace is a channel.

**Definition 2.18** (Choi Matrix)**.** Let $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ be a superoperator. Let $\{v_i\}$ be an orthonormal basis for $\mathcal{H}^A$ and $\gamma := (\sum_i v_i \otimes v_i)(\sum_i v_i^\dagger \otimes v_i^\dagger)$.[15] The Choi matrix of $\Phi$ is

$$J_\Phi := (I^A \otimes \Phi)(\gamma),$$

**Theorem 2.19** (Choi-Jamiołkowski Isomorphism)**.** *There exists a vector space isomorphism*

$$\begin{aligned} \mathbb{B}(\mathbb{B}(\mathcal{H}^A) : \mathbb{B}(\mathcal{H}^B)) &\xrightarrow{\quad\Delta\quad} \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B) \\ \Phi &\longmapsto J_\Phi. \end{aligned}$$

*The inverse isomorphism is*

$$\begin{aligned} \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B) &\xrightarrow{\quad\Delta^{-1}\quad} \mathbb{B}(\mathbb{B}(\mathcal{H}^A) : \mathbb{B}(\mathcal{H}^B)) \\ J &\longmapsto \big(\Phi : \rho \mapsto \mathrm{tr}_A((\rho^t \otimes I^B)(J))\big), \end{aligned}$$

*where $\rho^t$ is the transpose of $\rho$.*

*Proof.* Let $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ be a bounded superoperator, and let $\rho \in \mathbb{B}(\mathcal{H}^A)$. Via the given prescription, $\Delta(\Phi) = J_\Phi$, the Choi matrix. Then,

$$\begin{aligned} \mathrm{tr}_A((\rho^t \otimes I^B)(J_\Phi)) &= \mathrm{tr}_A\left(\sum_{ij} (\rho^t v_i v_j^\dagger \otimes \Phi(v_i v_j^\dagger))\right) \\ &= \sum_{ij} v_j^\dagger \rho^t v_i \Phi(v_i v_j^\dagger) \\ &= \sum_{ij} v_i^\dagger \rho v_j^\dagger \Phi(v_i v_j^\dagger) \\ &= \sum_{ij} \Phi(\rho_{ij} v_i v_j^\dagger) \\ &= \Phi\left(\sum_{ij} \rho_{ij} v_i v_j^\dagger\right) \\ &= \Phi(\rho), \end{aligned}$$

[15] This is the *scaled* maximally entangled state.

so $\Delta$ has an inverse on the left-hand side–notably, $\Delta^{-1}$–meaning it is an injection. On the other hand, suppose $J \in \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$ is an operator on the bipartite system. We wish to find a superoperator $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ so that $\Delta : \Phi \mapsto J$. Let $\{w_j\}$ be an orthonormal basis for $\mathcal{H}^B$. Let $\varphi \in \mathcal{H}^A \otimes \mathcal{H}^B$ be an arbitrary vector in the bipartite system. We can use our bases to write

$$\varphi = \sum_{ij} \rho_{ij} v_i \otimes w_j.$$

Then, define a function $r$ by

$$\begin{aligned} \mathcal{H}^A &\xrightarrow{\quad r \quad} \mathcal{H}^B \\ \varphi = \sum_{ij} \rho_{ij} v_i \otimes w_j &\longmapsto \sum_{ij} \rho_{ij} w_j v_i^\dagger. \end{aligned}$$

Let $\gamma_\ell$ be the vector $\gamma_\ell = (\sum_k v_k \otimes v_k) \in \mathcal{H}^A \otimes \mathcal{H}^A$, where $\gamma = \gamma_\ell \gamma_\ell^\dagger$. Then,

$$\begin{aligned} (I^A \otimes r)(\gamma_\ell) &= (I^A \otimes r)\left(\sum_k v_k \otimes v_k\right) \\ &= \sum_{ijk} v_k \rho_{ij} v_i^\dagger v_k w_j \\ &= \sum_{ij} \rho_{ij} v_i \otimes w_j \\ &= \varphi. \end{aligned}$$

Let $J \in \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$. We can always decompose $J$ via

$$J = \sum_i \zeta_i \eta_i^\dagger,$$

where $\zeta_i, \eta_i$ are all "simple tensors" of the form $u \otimes v$. In particular, we can find such $\zeta_i, \eta_i$ so that $\zeta_i \neq \eta_i$ for all $i$. Then, we may pick $r_i, \overline{r}_i : \mathcal{H}^A \to \mathcal{H}^B$ such that

$$\zeta_i = (I^A \otimes r_i)(\gamma_\ell)$$

and

$$\eta_i = (I^A \otimes \overline{r}_i)(\gamma_\ell).$$

Then, rewriting gives us

$$J = \sum_i \zeta_i \eta_i^\dagger = \sum_i (I^A \otimes r_i)\gamma(I^A \otimes \overline{r}_i)^\dagger.$$

Define a new superoperator

$$\begin{aligned} \mathbb{B}(\mathcal{H}^A) &\xrightarrow{\quad \Phi \quad} \mathbb{B}(\mathcal{H}^B) \\ (\cdot) &\longmapsto \sum_i r_i (\cdot) \overline{r}_i^\dagger. \end{aligned}$$

What happens when we plug $\Phi$ through $\Delta$? Well,

$$\begin{aligned}\Delta(\Phi) &= (\mathrm{id}^A \otimes \Phi)(\gamma)\\ &= \sum_{ij} v_i v_j^\dagger \otimes \Phi(v_i v_j^\dagger)\\ &= \sum_{ij} v_i v_j^\dagger \otimes \sum_s r_s v_i v_j^\dagger \overline{r}_s^\dagger\\ &= \sum_{ijs} (I^A \otimes r_s)(v_i \otimes v_i)(v_j^\dagger \otimes v_j^\dagger)(I^A \otimes \overline{r}_s)^\dagger\\ &= \sum_s (I^A \otimes r_s)\gamma(I^A \otimes \overline{r}_s)^\dagger\\ &= J.\end{aligned}$$

Thus, we may conclude that $\Delta$, is also a surjection. Together, we have shown that the given mapping $\Delta$ is a bijection. This is equivalent to stating that $\Delta$ is an isomorphism of the spaces $\mathbb{B}(\mathbb{B}(\mathcal{H}^A) : \mathbb{B}(\mathcal{H}^B)) \xrightarrow{\sim} \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$. □

The Choi-Jamiołkowsi isomorphism enables us to move between the channel picture and the bipartite operator picture freely. This is precisely the utility of a correspondence theorem. In particular, the Choi matrix $J_\Phi$ of a channel gives us useful information about channel characteristics of $\Phi$. We state these without proof, but they are good, simple exercises.

**Theorem 2.20** (Choi and Superoperators)**.** *If $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ is a superoperator, we can deduce some nice properties about $\Phi$ from looking at the Choi matrix $J_\Phi \in \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$:*

(*i*) $\Phi(\rho)^\dagger = \Phi(\rho^\dagger)$ *if and only if $J_\Phi$ is self-adjoint.*
(*ii*) *$\Phi$ is CP if and only if $J_\Phi$ is positive.*
(*iii*) *$\Phi$ is TP if and only if* $\mathrm{tr}_B(J_\Phi) = I^A$.
(*iv*) $\Phi : I^A \mapsto I^B$ *if and only if* $\mathrm{tr}_A(J_\Phi) = I^B$.[16]

## 2.2. **The Kraus Representation.**

**Theorem 2.21** (Operator-Sum Representation)**.** *A superoperator $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ is CP if and only if there exist Kraus operators $\{E_i : \mathcal{H}^A \to \mathcal{H}^B\}_{i=1}^r$ such that*

$$\Phi(\cdot) = \sum_i E_i(\cdot)E_i^\dagger.$$

*Proof.* Let $\{E_i : \mathcal{H}^A \to \mathcal{H}^B\}$ be a family of $r$ operators. For each $i \in \{1, \ldots, r\}$, the "conjugation" action $E_i(\cdot)E_i^\dagger$ is an isomorphism of $\mathcal{H}^A$. Thus, it is completely positive. Let $\rho \in \mathbb{B}(\mathcal{H}^A)$ be a positive operator. Then, $E_i \rho E_i^\dagger$ is also positive in $\mathbb{B}(\mathcal{H}^B)$. Thus, extending this to sums via the positive "cone" in $\mathbb{B}(\mathcal{H}^B)$, we see that $\sum_i E_i \rho E_i^\dagger \in \mathbb{B}(\mathcal{H}^B)$ is positive. Hence, $\Phi(\cdot) := \sum_i E_i(\cdot)E_i^\dagger$ is a completely positive superoperator. Conversely, let $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ be a completely positive superoperator. Then, per (ii), we know that the corresponding Choi matrix $J_\Phi$ is positive in $\mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$. Recalling the spectral theorem, we know we can write the Choi matrix in the form

$$J_\Phi = \sum_{i=1}^r \lambda_i \zeta_i \zeta_i^\dagger,$$

[16] That is, $\Phi$ is unital if and only if tracing out $A$ on the Choi matrix yields the identity on $B$.

and absorbing the coefficients $\lambda_i$ into the vector outer product, we can write

$$J_\Phi = \sum_{i=1}^{r} \eta_i \eta_i^\dagger.$$

Then, for each $i \in \{1, \ldots, r\}$, there exists an operator $E_i : \mathcal{H}^A \to \mathcal{H}^B$ such that $\eta_i = (I^A \otimes E_i)(\gamma_\ell) = \eta_i$. Let $\rho \in \mathbb{B}(\mathcal{H}^A)$ be an arbitrary operator on the $A$ system. Then, we certainly have

$$\Phi(\rho) = \mathrm{tr}_A((\rho^t \otimes I^B)(J_\Phi)) = \Delta^{-1}(J_\Phi),$$

where $\Delta^{-1}$ is the inverse map of the Choi-Jamiołkowsi isomorphism. We find

$$\begin{aligned}
\mathrm{tr}_A((\rho^t \otimes I^B)(J_\Phi)) &= \mathrm{tr}_A\left((\rho^t \otimes I^B)\left(\sum_{i=1}^{r}(I^A \otimes E_i)\gamma(I^A \otimes E_i)^\dagger\right)\right) \\
&= \sum_{ij} \mathrm{tr}_A\left((\rho^t v_j \otimes E_i v_j)(v_j^\dagger \otimes v_j^\dagger V_i^\dagger)\right) \\
&= \sum_{ijk}(\rho^t v_j v_k^\dagger \otimes E_i v_j v_k^\dagger E_i^\dagger) \\
&= \sum_{ijk} v_j^\dagger \rho v_k E_i v_j v_k^\dagger E_i^\dagger \\
&= \sum_i E_i \left(\sum_{jk} \rho_{jk} v_j v_k^\dagger\right) E_i^\dagger \\
&= \sum_i E_i \rho E_i^\dagger.
\end{aligned}$$

Hence, we may conclude that $\Phi(\rho) = \sum E_i \rho E_i^\dagger$, so the converse holds. Thus, we are done. □

*Remark* 2.22 (Naming). The Choi-Jamiołkowski isomorphism is also known as the "Choi isomorphism" or the "second canonical isomorphism." The operator-sum representation also goes by "operator-sum form" or "Kraus representation."

**Theorem 2.23** (Stinespring Dilation)**.** *Let $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ be a CP superoperator. There exists a Hilbert space $\mathcal{H}^C$ and an operator $E : \mathcal{H}^A \to \mathcal{H}^B \otimes \mathcal{H}^C$ so that*

$$\Phi(\cdot) = \mathrm{tr}_C\, E(\cdot)E^\dagger.$$

*Proof.* Suppose $\Phi : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$ be a completely positive superoperator. Via the Kraus representation, we can write

$$\Phi(\cdot) = \sum_{i=1}^{r} E_i(\cdot)E_i^\dagger$$

with Kraus operators $E_i : \mathcal{H}^A \to \mathcal{H}^B$ for all $i \in \{1, \ldots, r\}$. Consider the Hilbert space $\mathcal{H}^C := \mathbb{C}^r \simeq \mathbb{C}^{\oplus r}$ with orthonormal standard basis $\{e_i\}_{i=1}^r \subseteq \mathbb{C}^r$. Define an operator

$$\begin{aligned}
\mathcal{H}^A &\xrightarrow{\quad E \quad} \mathcal{H}^B \otimes \mathbb{C}^r \\
\zeta &\longmapsto \sum_{i=1}^{r} E_i \zeta \otimes e_i.
\end{aligned}$$

Define $\eta := \zeta\zeta^\dagger \in \mathbb{B}(\mathcal{H}^A)$. We compute

$$\begin{aligned}
\mathrm{tr}_{\mathbb{C}^r}(E\zeta E^\dagger) &= \mathrm{tr}_{\mathbb{C}^r}\left(\left(\sum_i E_i\zeta \otimes e_i\right)\left(\sum_i \zeta^\dagger E_i^\dagger \otimes e_i^\dagger\right)\right) \\
&= \mathrm{tr}_{\mathbb{C}^r}\left(\sum_{ij} E_i\eta E_j^\dagger \otimes e_i e_j^\dagger\right) \\
&= \sum_{ij} E_i\eta E_j^\dagger e_j^\dagger e_i \\
&= \sum_i E_i\eta E_i^\dagger \\
&= \Phi(\eta),
\end{aligned}$$

as $e_j^\dagger e_i = \delta_{ij}$; i.e., 0 if $i \neq j$ and 1 if $i = j$. Extending this linearly to a $\mathbb{C}$-linear combination, it is clear that the full statement of the theorem holds. □

2.3. **Anticliques and the Knill-Laflamme Condition.** Given a quantum error acting on our system, it would be desirable if we could determine its "correctability" via a condition on the error itself. First, we would like to know whether or not an error $\mathcal{E}$ is correctable in the context of our Hilbert space $\mathcal{H}$, performing so-called *error-syndrome measurements* to classify $\mathcal{E}$. Then, if $\mathcal{E}$ is correctable, we construct a *recovery operation* $\mathcal{R}$ to return the quantum system to its pre-error state. In order to do so, we make two major assumptions about the nature of the error:

(i) The error $\mathcal{E}$ is a quantum channel $\mathcal{E} : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})$.
(ii) The recovery $\mathcal{R}$ is another quantum channel $\mathcal{R} : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})$.

Now, even if much of $\mathcal{H}$ is "noisy" under the action of the error $\mathcal{E}$, we can look for a subspace of $\mathcal{H}$ to bury our information. We call such a subspace the *code space*.

**Definition 2.24** (Code Space)**.** A code space $\mathcal{C}$ is a $\mathbb{C}$-linear subspace $\mathcal{C} \subseteq \mathcal{H}$.

Then, given our information in the form of a quantum state $\rho \in \mathbb{B}(\mathcal{C})$, we call the map $\mathcal{R}$ a recovery operation on the code space $\mathcal{C}$ if

$$(\mathcal{R} \circ \mathcal{E})(\rho) \propto \rho,$$

where proportionality becomes precisely equality after the environment is traced out. The construction of a recovery operation $\mathcal{R}$ gives us a clear definition of correctability.

**Definition 2.25** (Correctable Error)**.** An error is correctable if there exists a code space $\mathcal{C}$ and such a recovery operation $\mathcal{R}$, defined as above.

**Definition 2.26** (Code Word)**.** The elements $v \in \mathcal{C} \subseteq \mathcal{H}$ which span the code space are called code words.

Given such a setup, we can now state in full the standard quantum correctability condition for QEC, often dubbed some variant of the "Knill-Laflamme subspace condition." The condition is both *sufficient* and *necessary* for a state $\rho \in \mathbb{B}(\mathcal{C})$ to be recoverable after an error was applied to the system.

**Theorem 2.27** (Knill-Laflamme)**.** *Let $\mathcal{E} : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})$ be a quantum error channel with Kraus operators $\{E_i\}_{i=1}^r$, and let $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$ be the orthogonal projection onto the code space $\mathcal{C} \subseteq \mathcal{H}$. Then, $\mathcal{E}$ is correctable if and only if*

$$PE_a^\dagger E_b P = \lambda_{ab} P,$$

*where $[\lambda_{ab}] \in \mathbb{M}_r(\mathbb{C})$ is self-adjoint.*

*Proof.* □

**Definition 2.28** (Error-Correcting Code)**.** Given a code space $\mathcal{C}$ and a set of correctable errors $\mathcal{E}$, the triple $(\mathcal{R}, \mathcal{E}, \mathcal{C})$ is a quantum error-correcting code.

*Remark* 2.29 (Anticliques). The orthogonal projection $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$ which "ignores" the non-code space part of the Hilbert space $\mathcal{C}^{\perp}$ is sometimes referred to as an *anticlique*.

**Definition 2.30** (Winter Space)**.** Let $\mathcal{E}$ be an error channel with Kraus operators $\{E_a\}_{a\in\Lambda}$. Then, the Winter space of the channel, first studied by Duan, is the space

$$\mathcal{V}_{\mathcal{E}} := \operatorname{span}\{E_a^{\dagger} E_b : a, b \in \Lambda\}.$$

Using this definition, we can rephrase the Knill-Laflamme condition with $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$ as

$$P\mathcal{V}_{\mathcal{E}}P = \mathbb{C}P,$$

saying $\mathcal{C}$ is a code space if and only if $\dim P\mathcal{V}P = 1$.

## 3. The Stabilizer Formalism

While the Knill-Laflamme subspace condition is elegant, it can practically difficult to work with. First, we will look at *noise* more carefully, attempting to understand the classical motivation for error correction. Then, we will take a look at correctability for errors based on the Pauli matrices. Finally, we will generalize our Pauli group to $n$ qubits, studying Gottesman's approach to stabilizer codes and the stabilizer formalism of quantum error correction.

### 3.1. An Elementary Code.

Let us consider the *binary symmetric channel* (Fig. 4). We send one classical bit betwen locations through a noisy classical channel. The channel noise causes the input bit to be flipped with probability $p > 0$, and with probability $1 - p$ the bit remains unchanged.

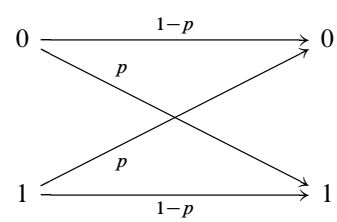


Figure 4. Binary symmetric channel

One simple way to combat this noise is to take three copies of each bit:

$$0 \to 000$$
$$1 \to 111,$$

where these new three-bit strings are called the *logical* 1 and 0. Sending one of our logical bit strings through the binary symmetric channel, suppose the output is 001. Then, assuming $p$ is not very large,[17] we may predict that the third bit was the one that was flipped, so the input to the channel was the logical 0. In general, as long as less than two of the bits are flipped, this form of *majority voting* decoding succeeds. It is then natural to ask: what is the probability of

[17] We prove the precise upper bound on $p$ soon.

two or more of the bits in our string flipping? Well, since one bit-flip error has probability $p$, we can compute the probability

$$\begin{aligned}\mathbb{P}(\text{two or more flips}) &= \underbrace{p^2(1-p)}_{\text{first two bits}} + \underbrace{p^2(1-p)}_{\text{first and last bits}} + \underbrace{p^2(1-p)}_{\text{last two bits}} + \underbrace{p^3}_{\text{all three bits}} \\ &= 3p^2(1-p) + p^3 \\ &= 3p^2 - 2p^3.\end{aligned}$$

Thus, given our encoding scheme $0 \to 000$ and $1 \to 111$, the probability of an error after majority voting decoding becomes $3p^2 - 2p^3$, whereas the probability of an error without any encoding was $p$. As such, the *repetition code* increases the reliability of our process when

$$3p^2 - 2p^3 < p \Rightarrow 3p - 2p^2 < 1 \Rightarrow p < \frac{1}{2},$$

using the quadratic formula.

3.2. **Stabilizer Groups, Spaces, and Codes.** It is common to rephrase correctability in the rich mathematical language of group theory. Recall the Pauli matrices

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{M}_2(\mathbb{C}).$$

Taking all possible products of these unitary, traceless $2 \times 2$ matrices, we were able to generate the *Pauli group*

$$\mathcal{P} := \langle X, Y, Z \rangle = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\},$$

of order/cardinality 16. The importance of the Pauli operators cannot be overstated when our setting is a 1-qubit space $\mathcal{H} \simeq \mathbb{C}^2$. The standard commutation and anti-commutation relations make Pauli operator computations far simpler. If we were instead to act on a multipartite system of $n$ qubits $\mathcal{H} \simeq \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = (\mathbb{C}^2)^{\otimes n}$, we would like a group which extends the multiplicative Pauli group to such $n$-fold tensor products. With this goal in mind, we may define the $n$-fold (or $n$-qubit) Pauli group

$$\mathcal{P}_n := \left\{ \gamma \bigotimes_{i=1}^{n} \sigma_i : \sigma_i \in \mathcal{P} \text{ and } \gamma \in \{\pm 1, \pm i\} \right\},$$

of order $4^n + 1$. As one would expect, the 1-qubit group $\mathcal{P}_1$ is precisely $\mathcal{P}$, as desired. Following common practice, we will write

$$X_j := I \otimes I \otimes \cdots \underbrace{\otimes X \otimes}_{j\text{th position}} \cdots \otimes I$$

for the 1-local action of the Pauli $X \in \mathcal{P}$ on the $j$th qubit. Likewise, $Y_j$ (resp. $Z_j$) is a 1-local application of the Pauli $Y$ (resp. $Z$) on qubit $1 \leq j \leq n$. As in the single qubit case, the $n$-fold Pauli operators exhibit useful commutation relations. Via the bilinearity of the tensor product, any sign changes from componentwise products become concatenated, so we can easily predict the behavior of products in $\mathcal{P}_n$ based on the number of parity swaps modulo 2.

*Remark* 3.1. Let $\{X_1, Z_1, \ldots, X_j, Z_j, \ldots, X_n, Z_n\}$ be a subset of $\mathcal{P}_n$ of cardinality $2n$. Then, we could consider an isomorphic copy (in the sense of sets and commutation relations) $\{\widehat{X}_1, \widehat{Z}_1, \ldots, \widehat{X}_j, \widehat{Z}_j, \ldots, \widehat{X}_n, \widehat{Z}_n\}$. These new operators can be thought of as acting on a collection of $n$ "virtual" qubits, rather than impacting the proper qubits of our space $(\mathbb{C}^2)^{\otimes n}$. Still, the $\widehat{X}_j, \widehat{Z}_j$, along with $\pm i$, generate all of our group $\mathcal{P}_n$.

Now, let $\mathcal{S} := \langle S_1, \ldots, S_s \rangle$ be an *abelian* subgroup of $\mathcal{P}_n$ such that $-I \notin \mathcal{S}$ and $s \leq n$. Without loss of any generality, we can assume that $S_j = \widehat{Z}_j$, our virtual qubit Pauli $Z$ operators, for all $j \in \{1, \ldots, s\}$. Certainly, the $\widehat{Z}_j$ commute with one another, ensuring $\mathcal{S}$ is abelian. Furthermore, this set is independent, in the usual sense for generators, so they can be diagonalized for all $j$, concurrently. Such a subgroup $\mathcal{S}$ will be called the *stabilizer*. We define a corresponding *stabilizer code space*

$$\mathcal{C} \equiv \mathcal{C}(\mathcal{S}) := \operatorname{span}_{\mathbb{C}} \left\{ |\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : \widehat{Z}_j \, |\psi\rangle = |\psi\rangle \text{ for all } 1 \leq j \leq s \right\}.$$

Now, $(\mathbb{C}^2)^{\otimes n} \simeq \mathbb{C}^{2^n}$, via the basis construction of the tensor product, so just by observation we can conclude that $\mathcal{C} \simeq \mathbb{C}^{2^{n-s}}$. If $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$ is the corresponding anticlique for the code space $\mathcal{C}(\mathcal{S})$, then $\widehat{Z}_j P = P$ for any $j \in \{1, \ldots, s\}$.

Now, recall the definition of the *normalizer* in a group $G$.

**Definition 3.2** (Normalizer)**.** The normalizer of $S \subseteq G$ is the subgroup

$$\mathcal{N}_G(S) := \{g \in G : gSg^{-1} = S\}.$$

On the other hand, we defined a closely related group known as the *centralizer*.

**Definition 3.3** (Centralizer)**.** The centralizer of $S \subseteq G$ is the subgroup

$$C_G(S) := \{g \in G : gsg^{-1} = s \text{ for all } s \in S\}.$$

Upon brief inspection, it should be evident that $C_G(S) \subseteq \mathcal{N}_G(S)$. Yet, in the case of our stabilizer subgroup $\mathcal{S} \subseteq \mathcal{P}_n$, we can say more: $\mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) = C_{\mathcal{P}_n}(\mathcal{S})$.[18] Using this group-theoretic language, we may state the major QEC stabilizer formalism theorem.

**Theorem 3.4.** *An error $\mathcal{E}$ with Kraus operators $\{E_i\}_{i=1}^r$ is correctable on a code space $\mathcal{C} = \mathcal{C}(\mathcal{S})$ if and only if*

$$E_a^\dagger E_b \in \operatorname{span}_{\mathbb{C}} \{\mathcal{P}_n \setminus \mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) \cup \mathcal{S}\}.$$

Intuitively, this result follows from the fact that $\mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) = C_{\mathcal{P}_n}(\mathcal{S})$ fixes the code space $\mathcal{C}$, as we would hope. Note that the normalizer takes on the form

$$\mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) = \langle i, \widehat{Z}_1, \ldots, \widehat{Z}_n, \widehat{X}_{s+1}, \ldots, \widehat{X}_n \rangle,$$

as evidenced by the commutation relations in $\mathcal{P}_n$.

## 4. Operator Quantum Error Correction

Thus far, our theory of error correction falls under the umbrella of quantum error correction (QEC). We now turn our attention to operator quantum error correction (OQEC), a formalism developed by Kribs, Laflamme, Poulin, and Lesosky in 2005.

[18] This useful identification follows from the commutation and anti-commutation of operators in $\mathcal{P}_n$. However, this is *certainly not* a general phenomenon for groups, so one must be careful when considering non-stabilizer subgroups in a similar context.

4.1. **Noiseless Subsystems.** Let $\mathcal{E} : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})$ be a quantum channel with Kraus operators $\{E_a\}$. Recall that an algebra is a set with addition, multiplication, and a $\mathbb{C}$-action. Additionally, recall that a C*-algebra, in finite dimension $n$, is just a subalgebra of $\mathbb{M}_n(\mathbb{C})$. Let $\mathcal{A} := \text{alg}\{E_a, E_a^\dagger\}$ be the finite dimensional C*-algebra generated by the Kraus operators $E_a$. Then, there is a unique decomposition (up to products by unitary operators) of the form

$$\mathcal{A} \simeq \bigoplus_J (\mathbb{M}_{m_J}(\mathbb{C}) \otimes I_{n_J}),$$

where $\mathbb{M}_{m_J}(\mathbb{C})$ is the square matrix algebra corresponding to $\mathbb{B}(\mathbb{C}^{m_J})$. Per usual, $I_{n_J} \in \mathbb{B}(\mathbb{C}^{n_J})$ is the identity on $\mathbb{C}^{n_J}$. We call $\mathcal{A}$ the *interaction algebra* of the channel $\mathcal{E}$.

**Definition 4.1** (Noise Commutant)**.** Given $\mathcal{A}$, as above, we define the noise commutant

$$\mathcal{A}' := \left\{\rho \in \mathbb{B}(\mathcal{H}) : E\rho = \rho E \text{ for all } E \in \{E_a, E_a^\dagger\}\right\}.$$

This is precisely all Kraus operators, including adjoints, which commute with operators in $\mathbb{B}(\mathcal{H})$. Thus, when the $\mathcal{E}$ error acts on a system, as long as $\mathcal{E} : I \mapsto I$, all of the states in the noise commutant $\mathcal{A}'$ are unaffected. Using the previous coproduct decomposition for $\mathcal{A}$, we see

$$\mathcal{A}' \simeq \bigoplus_J (I_{m_J} \otimes \mathbb{M}_{n_J}(\mathbb{C})),$$

up to products with unitaries. In particular, $\mathcal{E} : I \mapsto I$ implies that the noise commutant $\mathcal{A}'$ is precisely

$$\mathcal{A}' = \left\{\rho \in \mathbb{B}(\mathcal{H}) : \mathcal{E}(\rho) = \sum_a E_a \rho E_a^\dagger = \rho\right\} =: \text{Fix}(\mathcal{E}),$$

the *fixed set* of the channel $\mathcal{E}$. The usefulness of $\text{Fix}(\mathcal{E})$ is exactly why $\mathcal{A}'$ produces *noiseless subsystems* for $\mathcal{E}$. However, we cannot always assume that $\mathcal{E} : I \mapsto I$. We hope to find noiseless subsystems for $\mathcal{E}$ without relying on the structure of the noise commutant $\mathcal{A}$. Looking at the visual structure of the algebra $\mathcal{A}$, we can infer a nice decomposition for our Hilbert space

$$\mathcal{H} \simeq \bigoplus_J \mathcal{H}_J^A \otimes \mathcal{H}_J^B,$$

where we will call the $\mathcal{H}_J^A$ "noisy" subsystems with each $\mathcal{H}_J^A \simeq \mathbb{C}^{m_J}$, and call the $\mathcal{H}_J^B$ "noiseless" with each $\mathcal{H}_J^B \simeq \mathbb{C}^{n_J}$. Now, pull apart the decomposition:

$$\mathcal{H} \simeq \underbrace{(\mathcal{H}_1^A \otimes \mathcal{H}_1^B)}_{\text{Sector 1}} \oplus \underbrace{(\mathcal{H}_2^A \otimes \mathcal{H}_2^B)}_{\text{Sector 2}} \oplus \cdots \oplus \underbrace{(\mathcal{H}_J^A \otimes \mathcal{H}_J^B)}_{\text{Sector } J}.$$

Certainly, the simplest case is when our information is encoded within one "noiseless sector" of $\mathbb{B}(\mathcal{H})$, so let us simplify the decomposition to

$$\mathcal{H} = \underbrace{(\mathcal{H}^A \otimes \mathcal{H}^B)}_{\mathcal{C}} \oplus \mathcal{C}^\perp,$$

where $\mathcal{C}^\perp$ is just the remaining sectors summed together. Then, let $\dim \mathcal{H}^A := m$, $\dim \mathcal{H}^B := n$, and so $\dim \mathcal{C}^\perp := \dim \mathcal{H} - mn$.

4.2. **Revisitng the Stabilizer Formalism.** As we saw, our information is only encoded in the $\mathcal{H}^A$ subsystem of the code space $\mathcal{C} = \mathcal{H}^A \otimes \mathcal{H}^B$. Thus, two states are *logically equivalent* if they are equal on the $\mathcal{H}^A$ subsystem, ranging freely on the noisy $\mathcal{H}^B$ subsystem. We define a *gauge group* $\mathcal{G}$ to make precise this notion of state equivalence.

**Definition 4.2** (Equivalence Relation)**.** An equivalence relation $\sim$ on a set $S$ is a "binary relation"–that is, it takes two arguments in $S$ and compares them, in some sense–which for all $x, y, z \in S$, the relation $\sim$ follows

(i) *reflexivity*: $x \sim x$.
(ii) *symmetry*: $x \sim y$ implies $y \sim x$.
(iii) *transitivty*: $x \sim y$ and $y \sim z$ implies $x \sim z$.

**Example 4.3** (Classic Equivalences)**.**

(i) *Equality*: Consider equality $=$ on a set, say the integers $\mathbb{Q}$. Every number is equal to itself $a = a$. If $a = b$, then $b = a$. Finally, if $a = b$ and $b = c$, then $a = c$. The fact that equality is an equilvance relation is precisely the motivation for such a generalization.
(ii) *Isomorphism*: Consider isomorphism $\simeq$ on the collection of groups (or vector spaces). Then, we always have $G \simeq G$ using the identity homomorphism. If $G \simeq H$, then $H \simeq G$, using our inverse map. Finally, if $G \simeq H$ and $H \simeq K$, then $G \simeq K$ via the composition of isomorphisms.
(iii) *Similarity*: Consider similarity on the set of triangles in the plane $\mathbb{R}^2$. We have that the triangle $\triangle ABC$ is similar to itself. If $\triangle ABC$ is similar to $\triangle DEF$, then $\triangle DEF$ is similar to $\triangle ABC$. Finally, if $\triangle ABC$ is similar to $\triangle DEF$ and $\triangle DEF$ is similar to $\triangle HIJ$, then $\triangle ABC$ is similar to $\triangle HIJ$.

**Definition 4.4** (Quotient Group)**.** Let $(G, \cdot)$ be a group and $H \trianglelefteq G$ be a normal subgroup in $G$. We may form a quotient group $G/H$ which has elements in the form of "cosets" $gH$ for all $g \in G$, and an operation which is defined by $g_1 H \cdot g_2 H = (g_1 \cdot g_2)H$.

**Definition 4.5** (Gauge Group)**.** Take the quotient space of $\mathbb{B}(\mathcal{C})$ by the equivalence relation defined by $\rho \sim \rho'$ if and only if there exists a "gauge transformation" $g \in \mathcal{G}$ such that $\rho = g\rho' g^\dagger$.

In fact, the relation $\sim$ is an equivalence only if $\mathcal{G}$ satisfies the group axioms.

Now, we must have that $i \in \mathcal{G}$, as $\mathbb{B}(\mathcal{C})$ is conjugation-invariant for $i$. Similarly, the stabilizer group $\mathcal{S} \subseteq \mathcal{G}$ for precisely the same reason. Further, since the code space must be closed under the gauge equivalence, $\mathcal{G} \leq \mathcal{N}_{\mathcal{P}_n}(\mathcal{S})$. This implies that $\mathcal{G} \trianglelefteq \mathcal{N}_{\mathcal{P}_n}(\mathcal{G})$, so we can form the corresponding quotient group $\mathcal{L} := \mathcal{N}_{\mathcal{P}_n}(\mathcal{S})/\mathcal{G}$ of so-called "logical operations." Using our classification of the normalizer in the QEC case, along with the inclusions $\langle i \rangle, \mathcal{S} \subseteq \mathcal{G}$, we should be able to "complete" the gauge group with some virtual $\widehat{X}_j$ and $\widehat{Z}_j$ for $j > s$:

$$\mathcal{G} = \langle i, \widehat{Z}_1, \ldots, \widehat{Z}_s, \widehat{X}_{i_1}, \ldots, \widehat{X}_{i_a}, \widehat{Z}_{j_1}, \ldots, \widehat{Z}_{j_b} \rangle,$$

taking $\{i_k\}_{k=1}^a, \{j_k\}_{k=1}^b \subseteq \{s+1, \ldots, n\}$.

Via the Zanardi-Lidar-Lloyd axioms for inducing a tensor subsystem on the code space $\mathcal{C}$, the gauge and logical groups must admit a trivial commutator: $[\mathcal{G}, \mathcal{L}] = 0$. Thus, for $j \in \{s+1, \ldots, s+r\}$ and $s + r \leq n$, we must get pairwise $\widehat{X}_j$s and $\widehat{Z}_j$s:

$$\mathcal{G} = \langle i, \widehat{Z}_1, \ldots, \widehat{Z}_s, \widehat{X}_{s+1}, \widehat{Z}_{s+1}, \ldots, \widehat{X}_{s+r}, \widehat{Z}_{s+r} \rangle.$$

Taking the quotient, we get an isomorphic representation of the logical operators given by

$$\mathcal{L} \simeq \langle \widehat{X}_{s+r+1}, \widehat{Z}_{s+r+1}, \ldots, \widehat{X}_n, \widehat{Z}_n \rangle.$$

Via this identification, we will follow Poulin and interchange between $\mathcal{L} = \mathcal{N}_{\mathcal{P}_n}(\mathcal{S})/\mathcal{G}$ and the isomorphic representation via generators. Now, $[\mathcal{G}, \mathcal{L}] = 0$ via the Zanardi-Lidar-Lloyd axioms, and $\mathcal{G} \times \mathcal{L} \simeq \mathcal{N}_{\mathcal{P}_n}(\mathcal{S})$, so we get an induced subsystem structure on $\mathcal{C}$. This structure precisely states that for any logical operator $\ell \in \mathcal{L}$ and gauge operator $g \in \mathcal{G}$, restricted to $\mathcal{C}$ via the anticlique $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$, yields $gP = I^A \otimes g^B$ for some $g^B \in \mathbb{B}(\mathcal{H}^B)$, and $\ell P = \ell^A \otimes I^B$ for some $\ell^A \in \mathbb{B}(\mathcal{H}^A)$. Taking a look at dimensions, we get isomorphisms $\mathcal{H}^A \simeq (\mathbb{C}^2)^{\otimes k}$ and $\mathcal{H}^B \simeq (\mathbb{C}^2)^{\otimes r}$, of dimensions $2^k$ and $2^r$, respectively.[19] Thus, given $n = s + r + k$ virtual qubits arising from our isomorphic Pauli copies $\widehat{X}_j$ and $\widehat{Z}_j$, there are three ways to categorize:

(i) *s stabilizer* qubits: we have stabilizer generators $\widehat{Z}_j$ for $1 \leq j \leq s$, fixing the chosen code space $\mathcal{C} \simeq \mathbb{C}^{2^{r+k}}$.
(ii) *r gauge* qubits: we have $\widehat{Z}_{s+j}$ and $\widehat{X}_{s+j}$ generating the group $\mathcal{L}_B$ acting on non-encoding qubits. Such non-encoding qubits exist precisely to "absorb" any transformations from the gauge group $\mathcal{G}$.
(iii) *k logical* qubits: we have that the logical operations $\mathcal{L}$ are generated by $\widehat{Z}_{s+r+j}$ and $\widehat{X}_{s+r+j}$, acting solely on the $k$ encoding virtual qubits of $\mathcal{H}^B$.

*Remark* 4.6. Note that we have an isomorphism $\mathcal{G} \simeq \mathcal{L}_B \times \mathcal{S} \times \langle i \rangle$. Thus, if the gauge group $\mathcal{G}$ is abelian, then this approach precisely reduces to the classical QEC stabilizer formalism. This follows from the fact that the abelian case "erases" the product with the non-encoding group $\mathcal{L}_B$ and the product with $\langle i \rangle$, leaving $\mathcal{G} \simeq \mathcal{S}$.

Having built up the setting of QQEC and the gauge group $\mathcal{G}$, we can finally describe the OQEC correctability of noise channels $\mathcal{E}$ with Kraus operators $\{E_i\} \subseteq \mathcal{P}_n$. We will state and prove Poulin's stabilizer formalism for OQEC, which clearly reduces to the QEC case.

Measuring the stabilizer group generators $S_j = \widehat{Z}_j$ for $1 \leq j \leq s$, we will get a $s$-tuple of sign "outcomes" $(m_1, \ldots, m_s) \in \{\pm 1\}^s$. This $s$-tuple is the *error syndrome* of $\mathcal{E}$. If $(m_1, \ldots, m_s) = (1, \ldots, 1)$, then the state is in $\mathcal{C}$, the code space, whereas if $m_j \neq 1$ for any $1 \leq j \leq s$, then the error has shifted our state outside of $\mathcal{C}$. Using our knowledge of the commutation and anti-commutation relations in $\mathcal{P}_n$, we deduce that an error Kraus operator is "detectable" if it anti-commutes with one or more stabilizer generator $\widehat{Z}_j$, or if it fixes the encoded data.

**Theorem 4.7.** *Given an error channel $\mathcal{E}$ on $\mathcal{H} \simeq (\mathcal{H}^A \otimes \mathcal{H}^B) \oplus \mathcal{C}^\perp$ with Kraus operators $\{E_i\}_{i=1}^r$, a recovery channel $\mathcal{R}$ exists if and only if for all $1 \leq a, b \leq r$,*

$$E_a^\dagger E_b \in \operatorname{span}_{\mathbb{C}}\{\mathcal{P}_n \setminus \mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) \cup \mathcal{G}\}.$$

*Proof.* There are three possible cases, given any $a, b$:

(i) $E_a^\dagger E_b \in \mathcal{P}_n \setminus \mathcal{N}_{\mathcal{P}_n}(\mathcal{S})$.
(ii) $E_a^\dagger E_b \in \mathcal{N}_{\mathcal{P}_n}(\mathcal{S}) \setminus \mathcal{G}$.
(iii) $E_a^\dagger E_b \in \mathcal{G}$.

If (i), there exists $S \in \mathcal{S}$ so that the anti-commutator $\{E_a^\dagger E_b, S\} = 0$. Using the correctability condition with anticlique $P : \mathcal{H} \twoheadrightarrow \mathcal{C}$,

$$PE_a^\dagger E_b P = PE_a^\dagger E_b SP = -PSE_a^\dagger E_b P = -PE_a^\dagger E_b P = 0,$$

[19] The dimensions in [Pou05] are listed correctly in the isomorphisms, but are incorrect in the induced subsystem equations. We omit those dimension subscripts to avoid confusion.

so (i) yields correctable errors. If (ii), note that we have an isomorphism $\mathcal{N}_{\mathcal{P}_n}(\mathcal{S})\backslash\mathcal{G}(\mathcal{L}\backslash\{I\})\times\mathcal{G}$, so via the subsystem structure equations,

$$PE_a^\dagger E_b P = \ell_{ab}^A \otimes g_{ab}^B,$$

for some $\ell_{ab}^A \neq I^A$. Thus, (ii) yields errors which *cannot* be corrected. Finally, suppose (iii). Then, using the subsystem structure equations once more, we trivially satisfy

$$PE_a^\dagger E_b P = I^A \otimes g_{ab}^B,$$

so (iii) yields correctable errors.[20] □

Now, how would we go about recovering our information post-error? Well, if $E_a^\dagger$ and $E_b$ are equivalent errors, in the sense of $\mathcal{G}$, then

$$PE_a^\dagger E_b P = I^A \otimes g_{ab}^B,$$

thus yielding identical error syndromes.[21] Then, our measurement can further pair up each coset element $H \in \{E_a\}/\mathcal{G}$ with the corresponding error operator. Thus, take an arbitrary element of the coset $H$ and act on the state. Then, we would get an overall gauge transformation, thus preserving the $k$ logical qubits of $\mathcal{H}^A$.

## Acknowledgements

I thank Roy Araiza and Jihong Cai for their wonderful lectures on QEC, and OQEC, Eric Chitambar for giving me an intuition for channels, Anderson Trimm and Anastasia Perry for their input on this project, and Micah Fogel for his extensive guidance in facilitating an intersession of this kind.

[20] To state the theorem, we take the $\mathbb{C}$-linear span just to emphasize how the correctability can be extended to any complex linear combination. The result is equivalent.

[21] If $S \in \mathcal{S}$ and $g \in \mathcal{G}$, then the commutator $[gE_a^\dagger, S] = 0$ if and only if $[E_a^\dagger, S] = 0$. Thus, the gauge equivalence gives us identical syndromes.

## References


[ACC+24] Roy Araiza, Jihong Cai, Yushan Chen, Abraham Holtermann, Chieh Hsu, Tushar Mohan, Peixue Wu, and Zeyuan Yu. A note on the stabilizer formalism via noncommutative graphs. *Quantum Information Processing*, 2024.

[AL22] Roy Araiza and Felix Leditzky. Basics of finite-dimensional quantum information theory, 2022.

[Amo18] Grigori G Amosov. On general properties of non-commutative operator graphs. *Lobachevskii Journal of Mathematics*, 39:304–308, 2018.

[DF03] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Wiley, 3 edition, 2003.

[Dua09] Runyao Duan. Super-activation of zero-error capacity of noisy quantum channels, 2009.

[FIS21] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra*. Pearson, 5 edition, 2021.

[Got97] Daniel Gottesman. Stabilizer codes and quantum error correction, 1997.

[Hal17] Paul R. Halmos. *Naive Set Theory*. Dover, 2017.

[KLPL06] David W. Kribs, Raymond Laflamme, David Poulin, and Maia Lesosky. Operator quantum error correction, 2006.

[KLV00] Emanuel Knill, Raymond Laflamme, and Lorenza Viola. Theory of quantum error correction for general noise. *Physical Review Letters*, 84(11), 2000.

[NC16] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 10 edition, 2016.

[Pou05] David Poulin. Stabilizer formalism for operator quantum error correction. *Phys. Rev. Lett.*, 95:230504, Dec 2005.

[ZLL04] Paolo Zanardi, Daniel A. Lidar, and Seth Lloyd. Quantum Tensor Product Structures are Observable Induced. *Physical Review Letters*, 92:060402, February 2004.



University of Illinois Urbana-Champaign, Illinois, 61801

*Email address*: dheeran2@illinois.edu